0-10054

SR-E9-00404 NASA CR-160421
NAS9-15476

AgRISTARS

A Joint Program for Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing

Supporting Research

December 1979

(E80-10054) DEVELOPMENT OF LANDSAT-BASED TECHNOLOGY FOR CROP INVENTORIES Final Report, 15 Nov. 1978 - 14 Nov. 1979 (Environmental Research Inst. of Michigan) 217 p HC A10/MF A01 CSCL 02C G3/43 N80-18506 Unclas 00054

DEVELOPMENT OF LANDSAT-BASED TECHNOLOGY FOR CROP INVENTORIES

Q.A. Holmes, R. Horvath, R.C. Cicone, R.J. Kauth, W.A. Malila

ENVIRONMENTAL RESEARCH INSTITUTE OF MICHIGAN • P.O. BOX 8618 • ANN ARBOR, MICHIGAN 48107

NOTICES

Sponsorship. The work reported herein was conducted by the Environmental Research Institute of Michigan under Contract NAS9-15476 for the National Aeronautics & Space Administration, Johnson Space Center, Houston, Texas 77058. I. Dale Browne was Technical Monitor for NASA. Contracts and grants to the Institute for the support of sponsored research are administered through the Office of Contracts Administration.

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No. SR-E9-00404	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Development of Landsat-Based Technology for Crop Inventories		5. Report Date December 1979 6. Performing Organization Code
7. Author(s) Q.A. Holmes, R. Horvath, R.C. Cicone, R.J. Kauth, W.A. Malila		8. Performing Organization Report No. 132400-29-F 1
9. Performing Organization Name and Address Environmental Research Institute of Michigan Infrared and Optics Division P.O. Box 8618 Ann Arbor, Michigan 48107		10. Work Unit No 11. Contract or Grant No. NAS9-15476
12. Sponsoring Agency Name and Address NASA/Johnson Space Center Houston, Texas 77058 Attn: I. Dale Browne/SF3		13. Type of Report and Period Covered Final Report 15 November 1978 - 14 November 1979 14. Sponsoring Agency Code
15. Supplementary Notes The benefiting Supporting Research project element is Area Estimation Research.		

16. Abstract

This document describes research, development, testing and evaluation, and system design work undertaken at ERIM in support of on-going efforts of the Earth Observations Division of the NASA Johnson Space Center to apply aerospace remote sensing technology to agricultural inventory and crop condition assessment.

The general problem addressed is extraction of agronomic information at a large, if not global scale, of a type and quality that is relevant to decision makers, and in an efficient and cost-effective manner. A major undertaking was the development of concepts and goals toward which the technical effort related to information extraction could be directed.

This document addresses the general problem at three levels. First, conceptualization of the crop inventory problem and of possible area estimation approaches to it are presented. Second, technology issues and results are addressed in a format which partitions them into three complementary areas: labeling, machine processing, and performance evaluation. Finally, specific aspects of the technology are addressed in depth in 14 technical appendices.

17. Key Words Landsat, Agricultural Inventory, LACIE, Sampling, Performance Modeling, Performance Measures, Stratified Area Estimation, Labeling, Temporal-Spectral Features	18. Distribution Statement		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 207 + xxiv	22. Price

SR-E9-00404[1]
NAS9-15476

FINAL REPORT

DEVELOPMENT OF LANDSAT-BASED TECHNOLOGY FOR CROP INVENTORIES

BY

Q.A. Holmes, R. Horvath, R.C. Cicone, R.J. Kauth, W.A. Malila

The research reported here was initiated during the planning of the AgRISTARS Supporting Research Project and was a part of those plans, although this research will stand on its own merit. The benefiting Supporting Research project element is Area Estimation Research.

Environmental Research Institute of Michigan
P.O. Box 8618
Ann Arbor, Michigan 48107

December 1979

EXECUTIVE SUMMARY

This document describes research, development, testing and evaluation, and system design work undertaken at ERIM in support of on-going efforts of the Earth Observations Division of the NASA Johnson Space Center to apply aerospace remote sensing technology to agricultural inventory and crop condition assessment.

The research reported here was initiated during the planning of the AgRISTARS Supporting Research Project and was a part of those plans, although this research will stand on its own merit. The benefiting Supporting Research project element is Area Estimation Research.

The general problem addressed is extraction of agronomic information at a large, if not global scale, of a type and quality that is relevant to decision makers, and in an efficient and cost-effective manner. A major undertaking was the development of concepts and goals toward which the technical effort related to information extraction could be directed.

A three-stage approach was followed during the year:

1. Generically establish the basic design concepts and requirements of an information extraction system to provide the foundation for technology development.

2. Concentrate on the specific problem of area estimation and establish a baseline technology that provides an environment for technical growth and self-evaluation.

3. Pursue the research, development and testing of area estimation along two lines: that technology which relates to assigning crop labels to samples (called objective labeling), and that technology related to the efficiency of crop proportion estimation (called machine processing).

General Considerations

A perception of general requirements imposed by a global production and condition assessment system on supporting technology that employs remotely sensed satellite image data and collateral data is useful for focusing the technical efforts which are needed to support such a system. A general statement is that any information system ought to satisfy its users in the following general characteristics discussed in detail in Section 2:

Directivity: The ability to respond to the user's needs in a parametric sense.

Cost-Effectiveness: The ability to respond with sufficient information, in a timely manner, at a cost within the value of that information to the user.

Objectivity: The ability to provide the user with reliable estimates of the accuracy of the information.

When focused on the problem of crop area estimation based on remote sensing, these requirements for directivity, cost-effectiveness, and objectivity lead us to a baseline technology that acts as a framework for technology development. As discussed in Section 3, the technology is one based on stratified area estimation (SAE) and functionally consists of seven components: system tasking, data preparation, feature extraction, stratification, sample allocation, attribute assignment, and aggregation. Stratified area estimation can be implemented in a modular environment that permits directivity, can be analytically modeled so as to provide objectivity, and is a framework for phased development of crop acreage estimation technology that will lead to cost-effective component techniques.

Objective Techniques for Labeling

For the task of developing objective labeling techniques, an important effort was establisment of a four-step approach -- feature definition, feature extraction, signature characterization, and procedure development. A major goal is to achieve the best balance between manual and machine functions while developing objective techniques for labeling. The first year of a two-year effort is described in Section 4.

Using this approach, a refined machine algorithm for discriminating between spring wheat and barley using Landsat features was developed (predicated on having previously labeled spring small grains data):

- Segment-level features were established to indicate moisture stress and soil brightness.
- Feature extraction procedures were developed using temporal-spectral profile models.
- A discrimination rule was devised to adapt to differences in these indicator features.
- Testing is planned for first quarter of the next contract year.

Numerous other investigations were conducted to increase our understanding and/or capabilities in each of the major steps of the approach and to make progress toward the long-range goal. The accomplishments under each were as follow:

- Feature Definition
  - Relationships between the Landsat band ratios and the XSTAR-stabilized Tasseled-Cap plane were quantified.
  - A method for relating reflectance measurements to Landsat and Tasseled-Cap variables was defined.
  - Development of a meteorologically driven model of the spectral phenology of wheat was initiated.

- Feature Extraction
  - Products to facilitate analyst labeling of field-like blobs were developed.
  - Procedures and a mathematical theory for the production of color-stabilized image products were developed.
  - An improved spatial-spectral clustering algorithm was developed.
  - A study of crop development stage estimation was initiated.
  - Procedures for fitting profile model forms to multi-date Greenness values were developed for feature extraction and recommendations made for their use at several levels of application.
- Signature Characterization
  - Problems of extracting signatures and estimation with incomplete data were addressed; use of crop temporal-spectral profiles was suggested.
  - An analyst-labeling experiment was conducted to determine the pattern of analyst-labeling performance and attempt to characterize analyst-perceived signatures as a function of performance.

Implications of the reported investigations are discussed for the major sources of analyst-interpreter error that were identified in LACIE.

<u>Machine Processing Technology for Area Estimation</u>

The investigations of machine processing components for area estimation are described in Section 5 and performance evaluation studies undertaken in the context of Procedure M are presented in Section 6. These lead to statements we can make at three levels -- those related generically to overall area estimation technology, those related to

components of that technology, and those related to specific techniques employed as components.

In addressing issues related to the overall area estimation technology we find:

- Procedure M is a stratified area estimation environment that incorporates state-of-the-art technology in the areas of data preprocessing, feature extraction, stratification, sampling and aggregation; Procedure M utilizes a robust statistical framework, incorporates physical understanding of remote sensing of agronomic phenomena, and is implemented in a modular construction that enables monitoring of error propagation, enables analytic error modeling, and facilitates the comparative evaluation of existing and proposed area estimation component technologies.

With respect to major components of stratified area estimation technology we find:

- Data normalization including, at a minimum, sun angle correction, atmospheric correction, data screening and sensor calibration, is a crucial preprocessing stage that enables interpretation of data in a frame of reference in which agronomically related phenomena are relatively stable with respect to effects which impact signal value and are external to the crop phenomena.

- Stratification proves a useful tool in that it can lead to greater overall efficiency in producing estimates; major consideration should be given to the labeling interface in establishing optimum stratification approaches: the stratification of pixels into field-like shapes to provide optimum labeling targets, the stratification of these targets into pure-crop and mixed-crop quasi-fields to eliminate labeling errors associated with boundary, edge pixels and pixels in small fields, the stratification of data spectrally to produce homogeneous strata to which samples are directed, and the use of physically

based temporal-spectral strata that would permit, in time, estabishing prior expectations in each stratum with respect to crop content, crop condition, and labeling accuracy; though cost can be incurred in stratification of specific samples , if strata are not adequately homogeneous with respect to the crops being stratified, on the average a net cost is not expected as long as the stratification variables are positively correlated to crop type. Excessive labeling to supervise the stratification process can also lead to a net cost.

- The technique utilized for estimation of proportions of crops within a stratum should depend upon the nature of that stratum; greater error can be introduced by utilizing the wrong strategy than by ignoring more difficult strata.

- Evaluation of the performance of stratified area estimation entails the use of measures that describe efficiency of individual components in providing information to successive components and in propagating errors through the estimation system; stratified area estimation technology lends itself to error modeling that can provide insight into expected performance of the overall system, especially in the interaction of the error associated with labeled samples and its impact on system performance.

Related specifically to components and performance evaluation techniques utilized for area estimation in Procedure M, we find:

<u>For Components</u>

- Landsat 3 signals are attenuated in each band by 12 to 24% from corresponding Landsat 2 signals; however, an affine transformation has been defined that calibrates Landsat 3 to Landsat 2 which will permit the use of technology like XSTAR that was developed for Landsat 2 calibration.

- There is a cost/benefit to the use of spectral stratification; the benefit is derived in terms of sampling efficiency if homogeneous

strata are produced; a cost is incurred if the homogeneity is not sufficient to offset difficulty of allocating a fixed sample proportional to the size of the strata.

- Statistically based stratification techniques, like tolerance block or unsupervised clustering, are limited in their ability to consistently produce homogeneous strata of average purity greater than 85%. This limitation may be due to the features being used, the inherent limitations of separability of classes in MSS spectral space, or in the failure of the assumption that statistical distributions are directly correlated to crop classes.

- Stratification of pixels as field center and boundary before spectral stratification results in more homogeneous strata than combining the two, due to the confusion of mixed pixels and edge pixels as classes other than those contributing to their signal composition.

- Physically based stratification techniques, like the static spectral/temporal stratifier (SSTS), provide a low cost means for stratification that results in trajectory strata that are consistent from segment to segment, and appear comparable to statistically based strategies in terms of strata homogeneity.

- In order to minimize mean-square error:
  - Sampling strategies like Neyman or Bayesian sequential based on expected variation in crop proportion from stratum to stratum are made more feasible by the use of static stratification.
  - Labeling error introduces a variance into the system that should be considered when directing samples to strata.

- Bias in Procedure M can be reduced by simply sampling the stratum of little blobs; however, analyst labeling error in that stratum may introduce additional mean square error greater than that currently present.

- Non-parametric nearest neighbor classification approaches provide a possible mechanism for extending 'high confidence' analyst labels to other samples; this mechanism can produce 'probability labels' that depend on the spatial and spectral context of the sample.

For Performance Evaluation

- Through the use of analytical modeling, insight has been gained with respect to the impact of various components of a stratified area estimation procedure on the overall mean-square error of the system; the labeling component, in particular, has been modeled and shown to be significant in determining the system's capability of achieving a specific level of performance in terms of proportion estimation.

- Performance measures to be utilized in evaluating stratified area estimation procedures include sample purity, reduction of variance factors (especially the fixed-sample RV) which relate performance to classical statistical measures and, in addition, measures based on information theory; the latter are promising in that they naturally extend to multiple class problems.

As a supporting element for the investigations, a substantial data base of preprocessed Landsat data was prepared and is described in Section 7. Finally, recommendations based on the conducted investigations are summarized and assembled in Section 8. Together with this Executive Summary, they form a concise account of the year's effort and its ramifications.

PREFACE

This report describes part of a comprehensive and continuing program of research concerned with advancing the state-of-the-art in remote sensing of the environment from aircraft and satellites. The research is being carried out for NASA's Lyndon B. Johnson Space Center (JSC), Houston, Texas, by the Environmental Research Institute of Michigan (ERIM). The basic objective of this multidisciplinary program is to develop remote sensing as a practical tool to provide the planner and decision-maker with extensive information quickly and economically.

Timely information obtained by remote sensing can be important to such people as the farmer, the city planner, the conservationist, and others concerned with problems such as crop yield and disease, urban land studies and development, water pollution, and forest management. The scope of our program includes:

1. Extending the understanding of basic processes.
2. Discovering new applications, developing advanced remote-sensing systems, and improving automatic data processing to extract information in a useful form.
3. Assisting in data collection, processing, analysis, and ground-truth verification.

The research described herein was performed under NASA Contract NAS9-15476 and covers the period from November 15, 1978 through November 14, 1979. I. Dale Browne/SF3 was the NASA Contract Technical Monitor and Thomas Pendleton/SF3 was the primary NASA Technical Coordinator of the activity. The program was directed at ERIM by Richard R. Legault, Vice President and Head of the Infrared and Optics Division, Quentin A. Holmes, Program Manager, and Robert Horvath, Head of the Analysis Department.

The formally identified authors are those individuals who designed and coordinated both the research program and the preparation of this document. Actual authorship should recognize the contributions of the following individuals.

R.J. Balon	W.A. Malila
R.C. Cicone	W.F. Pont
E.P. Crist	D.P. Rice
H.M. Horwitz	W. Richardson
R.J. Kauth	T. Wessling
P.F. Lambeck	

The authors would like to acknowledge the research support provided by E.V. Lamson and D.J. Witte, and indicate their appreciation of the secretarial support provided by P.R. Swonger and D. Dickerson, with assistance from M.N. St.John, G.A. Goike and R.M. Coleman.

ERIM would like to extend its appreciation to others whose efforts facilitated the conduct of the program over the last year: R.L. Nance/SF3 of NASA/JSC for assistance in acquiring Landsat and ancillary data, and D.C. Helmer, C.V. Nazare, R.W. Payne and W.L. West of LEC/Houston for participation in the design and conduct of the analyst interpretation experiment.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont.)

TABLE OF CONTENTS (Cont.)

TABLE OF CONTENTS (Cont.)

TABLE OF CONTENTS (Cont.)

TABLE OF CONTENTS (Cont.)

(Appendices Contained in SR-E9-00404$_2$)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Cont.)

LIST OF TABLES

LIST OF TABLES (Cont.)

1
INTRODUCTION

Aerospace remote sensing technology has the potential to provide important contributions to agricultural inventory and assessment activities of the U.S. Department of Agriculture (USDA), other state and local agencies, and the private sector. Aerial photography has long had a role in operational activities of the USDA. As a result of the Large Area Crop Inventory Experiment (LACIE) [1], the USDA's Crop Condition Assessment Division of the Foreign Agricultural Service has developed a facility for routine use of Landsat data to monitor conditions in major crop production areas [2]. There still remains a substantial amount of research, development, testing and evaluation, and system design work in order to develop the technology to the point where its full potential can be realized.

The research reported here was initiated during the planning of the AgRISTARS Supporting Research Project and was a part of those plans, although this research will stand on its own merit. The benefiting Supporting Research project element is Area Estimation Research.

This report addresses several aspects of that development process and reports progress made in increased overall understanding and in advancing technology. Section 2 develops major system concepts for crop inventory and assessment. Then Section 3 further discusses area estimation technology which was the major focus of the contract work and introduces and provides a context for the details which are presented in the remaining sections, including objective techniques for labeling (Section 4), machine processing components for area estimation (Section 5), evaluation of area estimation performance (Section 6), description of a data base (Section 7), and recommendations (Section 8).

1.1 GENERAL CONTEXT

An agricultural inventory and assessment system belongs to a broad class of systems which could be used to affect, control, or monitor the environment. In very general terms, such environmental management systems consist of several parts--an information-gathering system, a forecasting system, a decision-making system, and an action-taking system--as discussed in Reference 3.

Briefly, an information-gathering system obtains data regarding the current state of the environment, including the results of past actions that affect the environment. A forecasting system requests and obtains information from the information system and, in view of a specific set of planned actions and a likely set of unplanned actions, produces an objective prediction of the future environmental state. The decision-making system hypothesizes a set of planned actions and obtains predictions of the resultant environmental state from the forecasting system. It decides among alternative sets of actions. The action-taking system carries out the planned actions and reports actions as they occur.

Because of the long lead times for technology development, it is natural to first develop the information-gathering component of an environmental management system, then the forecasting component, and last of all to create the possibility for coherent planned action by introducing a decision-making system. In developing the information and forecasting systems, it is wise to consider the characteristics that will be needed when operating in conjunction with a decision-making system. Notably, these are directivity, cost effectiveness, and objectivity.

By directivity, we mean an ability to respond to a user's changing information needs.

By cost effectiveness, we mean that the usefulness, accuracy, and timeliness of the estimates are commensurate with the cost of obtaining them. The distribution of system errors must be small enough that the outputs are useful for decision makers, both for regular predictions of a set of forecasted quantities and for special reports that may be required in a near-real-time mode. This emphasizes a need for a large, quality assured, data base, only a sample of which might be routinely accessed for regular scheduled forecasts.

Finally, by objectivity we mean, basically, believability. Some of the procedures which insure objectivity are that the forecasting process is visible to the decision makers in all essential elements, that the forecasts arise from fixed procedures applied to a data base, that the data base be subject to a rigorous quality assurance procedure, that the actual quantities forecasted are quantities that will subsequently be known with accuracy significantly better than the forecast accuracy, that the system publishes its estimated error distribution along with its forecasts, and that the system publishes posterior comparisons of its forecasts with the subsequently known forecasted quantities.

A summary of types of agricultural information that are potentially extractable from aerospace remote sensing data is presented in Table 1. The first is crop identification which has received a majority of the attention in agricultural studies to date, especially in conjunction with crop area estimation. Next are indications of crop development stage and crop condition which could provide important inputs to yield models. Soils are a topic that have not received very much emphasis to date, but they too have an important effect on yield and productivity. Together, estimates of crop area and crop yield permit estimate of overall crop production, the "bottom line" of agricultural crop inventories.

TABLE 1. POTENTIAL CONTRIBUTIONS OF AEROSPACE REMOTE SENSING TO AGRICULTURAL INVENTORY AND ASSESSMENT

- Crop Identification
- Crop Development Stage
  - Planting and Harvesting Progress
  - Key Growth/Development Stages
- Crop Condition
  - Vigor, Stress
  - Ground Cover, LAI
  - Management Practices
  - Homogeneity
  - Episodal Events
- Inputs to Yield Models
  - Spectral
  - Meteorological
- Soil Characteristics
- Crop Area
  - Total Area Planted, Harvested
  - Area by Condition Class
- Crop Production

In LACIE, remote sensing provided identification of small grains and subsequent estimates of their planted area. First-order spectral indicators of crop stress were also developed. During LACIE transition years 1978 and 1979, some research was directed at first-order spectral indicators of crop development and improved spectral indicators of crop stress. Use of these spectral characteristics to improve labeling and adaptation of the technology from wheat and small grains to corn and soybeans was also initiated.

The full potential of remote sensing contributions to agricultural inventory and assessment has not yet been realized. There will be improvements in sensors (e.g., thematic mapper and meteorological satellites), information extraction techniques, inventory system technology, and in joint use of meteorological and spectral data. In the program planned for AgRISTARS (Agricultural and Resources Inventory Survey Through Aerospace Remote Sensing), this technology will be developed and evaluated in additional geographic regions and for additional crops.

2
CROP INVENTORY AND CONDITION ASSESSMENT CONCEPTS

This section provides an overall context for Section 3 and following sections which describe the past year's work on the major area estimation components of a crop production estimation system. It discusses general requirements on large-scale, perhaps global production and condition assessment systems based on remotely sensed satellite image data and collateral data; it also indicates our perception of the technical efforts which are needed to bring such systems into being. In describing these concepts, we are not excluding the possibility that a particular user may wish to focus on some specific system aspects or route of development. Rather, we are attempting to draw out conclusions that will be valid in a generic sense.

2.1 GENERAL SYSTEM CHARACTERISTICS

Any information system ought to satisfy its users in the three general characteristics mentioned earlier: directivity, cost effectiveness, and objectivity. Directivity means that the system has a range of functions satisfactory to the user's needs and that he can ask for any of those functions to be executed and receive information back in a timely manner.

Cost effectiveness means that a user on the average receives information more valuable than the cost of producing it and that the cost is lower than the cost of the same information from a competing source. In this context, more accurate information is generally more valuable, and the user has to take accuracy into account in computing value.

Objectivity means that the system accurately states its error bounds, i.e., provides reliable estimates of its own accuracy. Without this characteristic, the user cannot determine the value of the information he receives, and in fact cannot use the estimates in a rational manner.

These generic attributes, taken in context with the technological possibilities we can see over the next ten years, imply to us some more specific system characteristics and high priority development issues. Table 2 summarizes the detailed discussion which follows.

2.1.1 DIRECTIVITY

Directivity implies a command structure to the system. The various likely system functions will rest upon the use of common system components, used in a different way for each application. For example, qualitative condition assessment may use imagery and collateral weather data in a mode in which certain general characteristics (e.g., Greenness) are extracted automatically while an analyst identifies other characteristics (e.g., low reservoirs). The same imagery, collateral data, and analyst comments might be used as priors in an area estimation subsystem. As another example, the same subsystem that lays on samples for producing area estimates may be used to lay on samples for extracting spectral inputs to a yield estimate. Thus directivity implies a command structure which can link together various subsystems in various configurations to respond to a variety of user demands.

2.1.2 COST EFFECTIVENESS

Topics under this heading include sampling, measurement accuracy, and timeliness.

The system requirement which most obviously arises from cost consideration is the requirement for sampling. We do not expect that systems arising from the current and next generation sensor and data handling technology (i.e., systems operational in the 1985 to 1995 time frame) can be cost effective without sampling being an integral part of system design and operation. The subject of sampling may be organized in various ways as indicated in Table 3. Each "cut" provides some different insight into the application in remote sensing based systems.

TABLE 2. INVENTORY AND CONDITION ASSESSMENT LONG RANGE GOALS AND ESSENTIAL ELEMENTS

- The long-term goal is a high degree of:
  - directivity
  - cost effectiveness
  - objectivity

  consistent with the basic information sources available at a particular point in time.

- Jointly, these goals imply a modular system whose essential elements are:
  - command structure
  - sampling
  - measurement
  - aggregation
  - accuracy assessment

  operating upon

  - data base derived from currently available sensors

  within an environment provided by

  - supporting systems

TABLE 3. DEVELOPMENT OF SAMPLING VIEWED FROM VARIOUS ASPECTS

- Development and Evaluation of
  - Stratified sampling
  - Multistage sampling
  - Sequential sampling
- Incorporation of Auxiliary Information
  - Prior year estimates
  - Current year weather and other collateral information
- Multiple Aspect Sampling and Aggregation
  - Area, yield, production
  - Multiple crops
  - Multiyear
  - Multisensor
- Flexible Sampling Strategies
  - Full frame based
  - Automatic resampling when samples are lost due to clouds, etc.
  - Automatic resampling when self assessment indicates need for more samples to meet an accuracy goal

Sampling may be appropriate at any level (local/global) in the system. In general, wherever sampling is employed the pattern is the same:

Sampling (select units from a sample frame)
Measurement (for selected units)
Aggregation (estimate for sample frame)

The nesting of this pattern from the global to the local condition constitutes hierarchical or multistage sampling schemes as illustrated in Figure 1.

The age or timeliness of an estimate also affects the value of information and hence cost effectiveness. The need for a timely estimate suggests sequential sampling schemes in which each sequential stage relies on information obtained in the previous stage to establish how many additional samples are needed to converge upon an estimate within given error bounds. Such estimates would not only be achieved with minimal sampling, they would also be produced in minimal time. At every stage, hierarchical or multistage sampling schemes may be unbiased with respect to the estimates from the next lowest stage. Estimation error at the final stage drives the errors of the entire system. It is not conceptually reasonable to attempt to design an unbiased sampling scheme which in any way "corrects for" or is "robust against" measurement errors whose statistical properties are unknown.

Hence the measurement errors at the final (most local) point in the system are a critical element in cost effectiveness. If there are systematic errors in measurement, no amount of sampling will, by itself, make the derived information unbiased and thus of more value. If there are large random errors in measurement, the sampling cost must be increased to obtain more valuable estimates; but more importantly, the number of contol (ground truth) samples must be large in order to place bounds on the unknown bias of the measurements. (Without such bounds, the system is not objective in the sense defined here.) Hence, measurement accuracy is identified as a single highest priority issue in system development. This is reflected in

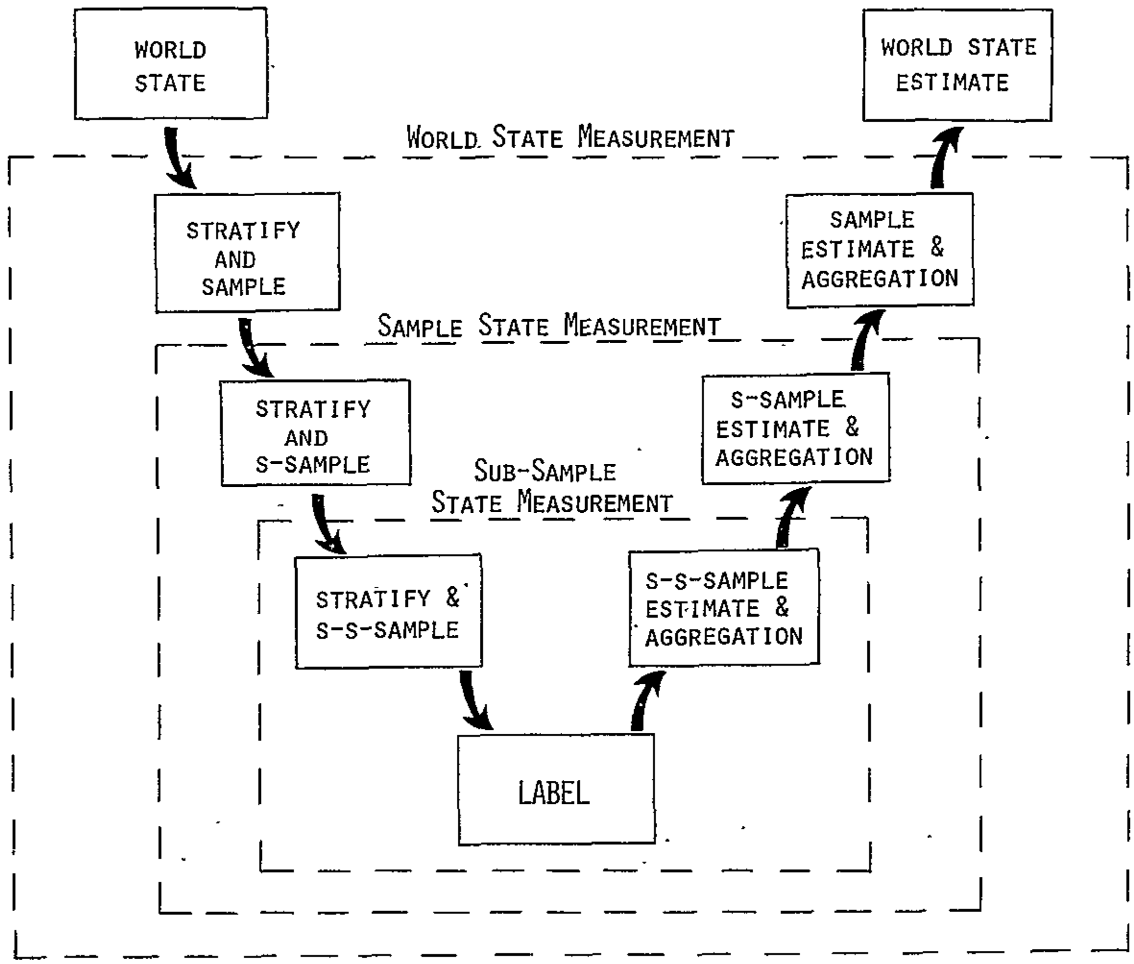


FIGURE 1. NESTING PATTERN OF MULTISTAGE SAMPLING

Section 3 and following in which the problem of objective labeling is given a dominant role. Table 4 summarizes issues of measurement accuracy for area estimation which will appear in discussion throughout this report.

2.1.3 OBJECTIVITY

Objectivity is measured by the honesty with which a system identifies its errors. Objectivity is only achieved through a thorough assessment of the sources of error in a system, i.e., through a validated system performance model which allows the statement of the estimated error distribution associated with each primary estimate. If the primary estimate is a vector (e.g., area or production estimates for several crops), then the estimated joint error distribution should be stated (e.g., the covariance of the crop estimates). The performance model should be modular in the same sense that the system of making estimates is modular, so that separate performance estimates can be made for each component and for the system as a whole. Component performance models can then be validated on a component-by-component basis during system development.

Component and system performance models have two important additional roles during system development, as will be discussed later.

2.2 DEVELOPMENT CONCEPTS

Given overall system criteria and a general system concept which includes major system components, the problem of detailed definition of components arises. While the major component structure may appear reasonably straightforward, the detailed definition involves consideration of a large number of possibilities for each major component, as is suggested by Tables 2 and 3 which outline some of the possibilities for area estimation and sampling, respectively.

TABLE 4. SCOPE OF REQUIRED DEVELOPMENTS IN MEASUREMENT ACCURACY

- Preprocessing Data Normalization
- Feature Extraction
  - Physically based through models and supporting field research
  - Statistically based through research in classification
- Estimation Techniques
  - Development of classifier technology
  - Development of better analyst/machine interfaces
  - Development of direct proportion estimation techniques
  - Development of signature characterization techniques
  - Association of collateral information in any of the above

How will we choose among the multiple possibilities for system development? This is the challenge of controlled development. It seems to us that a useful approach is to seek stepwise improvement for each component by developing and evaluating alternatives. Periodically the developed components ought to be brought together and evaluated as a system. This could be accomplished through a series of overlapping Technology Phases as shown in Figure 2, each encompassing directed research, component development, component evaluation, system integration and system evaluation. In this approach there will be opportunity for the accumulation of experience in the technical disciplines associated with each component. System integration and pilot type demonstrations periodically provide beneficial exposure to the harsher aspects of reality. (Such an approach is implicit in the AgRISTARS plan for the Foreign Commodity Production Forecasting Project which envisions an overlapping series of pilot tests which differ in the countries of application, the crops of interest, and the technology and sensor level to be used.)

In order to keep track of what has actually been learned and accomplished, and in order to evaluate and choose between alternative components during development, a system and component performance model is essential. Each developed component ought to have an appropriate performance model associated with it; and it ought to be the responsibility of the component developer to provide the form of that model (since the developer has, presumably, the best insight into the expected behavior and error sources of the component). Test and evaluation ought to focus both on evaluating a component and validating the performance model for that component.

Each component may have several performance measures, appropriate to the type and the level of that component in an overall system. The evaluation of a component ultimately is with respect to its marginal impact on system performance. Hence, early in the development cycle

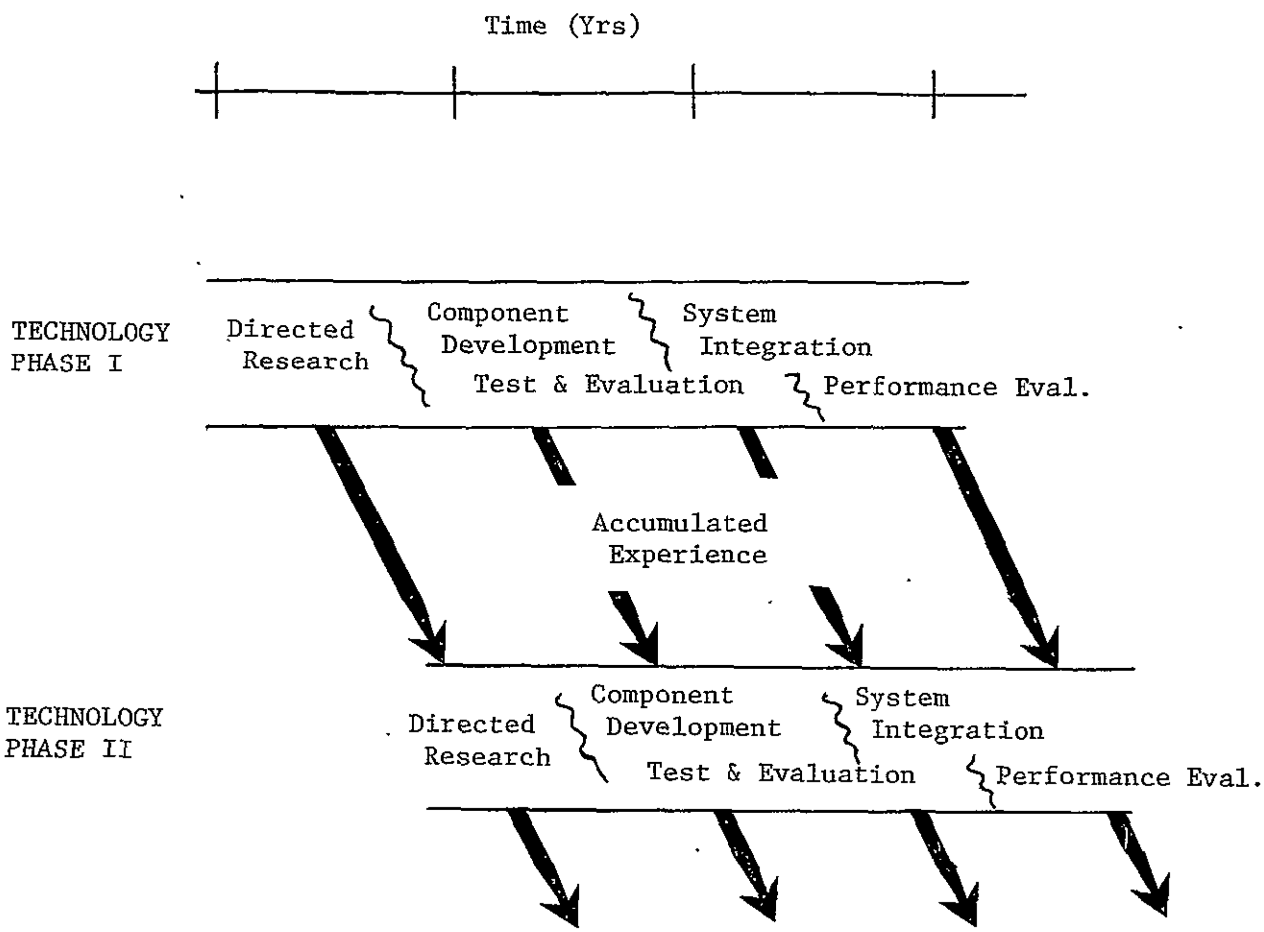


FIGURE 2. ILLUSTRATING THE CONCEPT OF OVERLAPPING TECHNOLOGY PHASES AND THE ACCUMULATION OF EXPERIENCE

there needs to be a baseline system performance model, with all components in place, in order to evaluate development alternatives for any single component.

The actual conduct of a complex development strategy such as that indicated above would place heavy demands on the underlying data management and computational support system. We will not discuss this topic, but we note that emphasis in computation should be on a modular, flexible, processing system which yet has the capability of processing relatively large amounts of data in order to conduct system-scale tests. ERIM has developed the QLINE processing system which combines these desirable features.

The above discussion is summarized in the following statements:

- Considerations about learning and accumulation of experience suggest a series of system implementations (technology phases), each one building on aspects of the previous ones.
- Within a technology phase, certain types of development may duplicate functions or conflict with each other -- hence they must be evaluated as alternatives.
- With limited development resources, priorities must be set to establish which of the multiple possible paths of growth will actually be followed.
- Component technologies can grow and develop independently -- but their state of development can only be evaluated in a system context.
- Performance modeling and configuration control are essential organizational elements to maintain understanding of what is developed.
- All system elements must be present in embryo in first implementation.
- A flexible modular computational and data management system is needed to support multiple and changing implementations.

2.3 TECHNICAL ISSUES

For three of the technical topics noted in the above discussion (Sections 2.1 and 2.2), we have identified technical issues. These topics are: area estimation, sampling, and performance modeling.

2.3.1 AREA ESTIMATION

Key technical issues which arise from consideration of the area estimation problem are:

- Multitemporal image registration
- Signature characterization
- Mixed and boundary pixels
- Analyst/machine interfaces
- Error characterization for both analyst and machine processes.

All of these topics except multitemporal registration are discussed in later sections of this report. Here we will say a few words about multitemporal registration.

Accurate multitemporal registration will be essential in a final operational system, since multitemporal image data are required in order to obtain reasonable crop classification accuracy with Landsat data. However, the accuracy on multitemporal registration need not be as stringent during development activities. For instance, data from field centers should be adequate for the development of signature characterization technology. Similarly, while the final solution of the mixed and boundary pixel problem will require accurate multitemporal registration, initial developments might proceed with unitemporal data or data registered with current levels of accuracy. Hence multitemporal registration can develop for some time as an independent topic.

Considerations of the need for flexible sequential sampling strategies suggest a three-step approach for acquisition of multitemporal registered images (sample segments). In Step 1, a full-frame fitting procedure would be carried out and the coefficients of the fit would be stored. In Step 2, the raw Landsat data (A-Tapes) would be accessed in response to a data request. The selected multitemporal set of images would be closely (though not precisely) matched by utilizing knowledge of the fitting coefficients developed in Step 1. The images would then be screened and a decision made as to their utility. If a sample segment passed the screening tests, then Step 3 would take place. In Step 3, the image data would be used to make local adjustments to the registration coefficients, the data would be spatially resampled, and the desired set of registered multitemporal images would be produced. The major advantage of this approach is that segment selection could be carried out prior to the highly accurate (and costly) final registration and resampling.

2.3.2 SAMPLING

Key technical issues which arise from consideration of the sampling problem are

- Multiyear sampling
- Sequential sampling (both at the global and within segment levels)
- Use of collateral and analyst inputs to the stratification process

This report discusses the second and third items at the segment level. On-going and planned efforts by NASA and its supporting research contractors deal with the first two issues at broader levels.

2.3.3 PERFORMANCE MODELING

Key technical issues which arise from considerations of performance modeling are:

- System performance
- Sampling/aggregation component performance
- Measurement performance

Sampling/measurement/aggregation are discussed in this report at the segment level only in Section 6.1.

3
AREA ESTIMATION TECHNOLOGY

Area estimation is the component of overall crop inventory technology that is the focus of investigation in this report. In this section, we present several prevalent approaches to crop area estimation based upon remote sensing, emphasizing a generic description of stratified area estimation technology. Then, the related key technical issues addressed in this report are introduced.

Common to the approaches is a need for some type of representation of the classes of interest. In the most general case, the area estimation technology has knowledge of ground cover or crop vegetation canopy 'signature Signatures are ultimately a functional description of observed features that are unique to, or at least descriptive of, the canopies under varying conditions. Often the term 'signature' has been used in a restrictive sense to describe the means and covariance matrix computed from labeled spectral observations of a canopy. Surely these are estimators of true signatures under the assumption that spectral data from crop canopies are either normally distributed statistically or can be effectively described as sums of normal distributions; complete signatures must include collateral variables as well. Crop area estimation approaches can also use attributes or 'labels' assigned to samples drawn from the area of interest.

We shall discuss, in this section, a number of machine-oriented techniques to utilize labeled samples or machine signatures to produce estimates of the areas of specific canopies of interest.

3.1 MACHINE-ORIENTED APPROACHES TO CROP AREA ESTIMATION

A most accurate way to estimate the ground area from which a particular crop of interest will be harvested would require selecting a principle sampling unit, say the farmer's field, and exhaustively labeling each field from ground observations. Some registered remotely sensed data

source, like aerial photography or a multispectral sensor, could be used for purposes of efficiency to aggregate the area of each crop of interest. Certainly, when one's concern is for estimates over large areas, even on the global scale, the approach of total enumeration is not feasible. Hence, a number of approaches are utilized, usually with sampling, to reduce the costs of inventory systems to be within budgetary constraints and at the same time maintain an accuracy that is acceptable. Three approaches to making estimates are: (1) classifiers, (2) label-based stratified area estimators, and (3) signature-based proportion estimators. A brief discussion of each approach ensues.

3.1.1 CLASSIFIERS FOR AREA ESTIMATION WITH REMOTE SENSING

A traditional approach to area estimation with remote sensing is to utilize classifiers prevalent in the literature of pattern recognition. The basic process is as follows:

(1) Select a sample set that is designated 'training'.

(2) Based on this sample, form a functional description, usually Gaussian, of each crop of interest.

(3) Select a decision rule or discriminant that would best separate the classes of interest.

(4) Use the sample to train the discriminant and classify all sensor signals with it.

(5) Aggregate the result to produce estimates.

Early in the application of this approach, discriminants based on quadratic or simple linear rules were applied to spectral data. Often promising results were achieved; especially when the training set or procedures could be tuned to the need. An early experience of the large-scale application of this approach was the Crop Inventory Technology Assessment from Remote Sensing (CITARS) experiment [4]. It was learned that, when subjected to the rigor of standardized requirements. classifiaction approaches lacked consistency and did not seem to live up to the

promise expected of them. Early in the Large Area Crop Inventory Experiment (LACIE), the problem of defining adequate training for a classifier became apparent. The need for estimates with bias and variance within measurable tolerances of error led to emphasis on the statistical rigor of the estimation procedure. This resulted in the stratified area estimation approach embodied in Procedure 1 [1].

This is not to say that estimation approaches based on classifiers are without merit. Indeed, the efficiency associated with a relatively small training requirement and the fact that classification of each signal reduces sampling error to a negligible amount is desirable. Research in the area of classifiers has begun to uncover new potential. A recent study by S. Wheeler [5] has illustrated a means of estimating the bias associated with a simple Fisher linear discriminant. Work by H. Horwitz, reported in Appendix N, illustrates a similar result using a nonparametric classifier. In addition, approaches utilizing spatial information are being developed, as well as strategies utilizing multitemporal measurements in unique ways.

3.1.2 STRATIFIED AREA ESTIMATIONS BASED ON LABELED SAMPLES

An effort to lend statistical rigor to crop area estimation in LACIE led to the development of Procedure 1 [1], a stratified area estimation (SAE) approach used within sample segments. The SAE approach was designed to provide estimates targeted to stated precisions in a consistent manner. More will be said in Section 3.2 about SAE, since it forms a focus of activities reported herein.

Stratified area estimation is based on statistical theory utilizing stratified random sampling as a framework for estimation of crop acreage. Samples are allocated, labeled, and estimates therefrom aggregated. The sample size and the accuracy with which they are labeled will drive the precision and bias of the system.

3.1.3 PROPORTION ESTIMATORS BASED ON SIGNATURES

A unique approach to the estimation of crop acreage is based on utilizing crop signatures to produce estimates directly of crop proportion as a function of the underlying distributions of spectral data.

Two approaches under development include UHMLE [6] and CLASSY [7]. The UHMLE proportion estimator is an iterative approach utilizing maximum likelihood techniques and signatures based on training data. With each iteration of the procedure, prior probabilities associated with Gaussian representations of the crop signatures are updated. It is expected that successive estimates of these priors converge to a stable result. Sensitivity of the procedure to initial signatures is a major concern. Current studies [8] are examining use of spatial features to establish more accurate signature representations.

CLASSY extends the UHMLE concept by incorporating a statistical representation of distributions that incorporates four moments. A sophisticated split/combine algorithm covers the underlying distributions in a manner unmatched by any other approach.

It remains to be seen whether sophisticated statistical approaches can uncover distributions that are truly correlated with classes of interest, and whether the need for accurate signature representation can be adequately met.

3.2 STRATIFIED AREA ESTIMATORS

The use of stratified area estimators for crop acreage inventory was initiated in LACIE with the development of Procedure 1. Major developments in the technology to date can be represented by three phases:

(1) The introduction of a robust statistical framework to crop acreage inventory.

(2) The incorporation of physical understanding of phenomena related to crops, crop practices, and remote sensor conditions.

(3) Developments toward greater efficiency in terms of precision and cost.

With Procedure 1, one first finds a concern with sampling as a means to provide estimates at a given precision and to identify sources of error.

The development at ERIM of Procedure B [9] in the same time frame as Procedure 1, and later Procedure M [3], introduced, to a statistical framework closely related to that of Procedure 1, a physical understanding of (1) sensor response to atmospheric and solar conditions (the use of haze and sun angle normalization), (2) sensor response to reflectance phenomena (the use of the Tasseled Cap and other multitemporally-based features), (3) agronomic cropping practices (the use of fields as basic sampling units and determination of approximate planting dates), and (4) crop phenomenology and growing conditions (the use of green development trajectories and the effects of stress and soil color upon those trajectories). These elements are described later in this report.

The third area of development is in greater efficiency in terms of precision and cost. Advanced developments in alternative procedures based on Procedure 1 have resulted in the use of advanced sampling strategies that require fewer samples for a given level of efficiency. A multisegment approach, wherein segments are grouped and samples are selected from a subset of segments, is an optional capability incorporated within Procedure M.

Stratified area estimation approaches to crop inventory provide an excellent baseline from which area estimation technology as a whole can grow. An approach well founded on statistical principles, such as SAE, permits a tracking of errors that was found difficult in more traditional classifier approaches. One of the important results of this capability

was a confirmation of the importance of accurate labels and the complexities associated with the labeling process. In the next two subsections, we shall briefly describe the basic components of a stratified area estimation strategy and issues related to its use.

3.2.1 COMPONENTS OF AN SAE TECHNOLOGY

In the view of this report, stratified estimates of crop acreage within a sample segment are functionally structured in the following manner:

- System Tasking: A specific configuration is defined to respond to a specific set of information needs, i.e., the SAE technology is parameterized.
- Data Preparation: A resource bank of data is formed and preprocessed in a manner that facilitates information extraction.
- Feature Extraction: Information relevant to the need is extracted from the data.
- Stratification: Data are grouped into strata that are homogeneous with respect to the crop of interest.
- Sample Allocation and Selection: Acreage estimates of each stratum will be made based on samples drawn in some efficient, unbiased manner.
- Attribute Assignment: Sampled data entities are described in a meaningful way, such as labels of crop type and/or condition.
- Aggregation: The assigned attributes are used to generate stratum-level area estimates that in turn are aggregated to provide segment-level estimates.

Procedures 1, 1A, and M can all be described in the context of this generic definition of stratified area estimation approaches. This definition will prove useful in organizing and addressing key issues of this report.

3.2.2 ISSUES RELATED TO SAE

Four issues are the key to the continued successful development of SAE technology. These are: (1) improved techniques for attribute assignment, i.e., labeling, (2) greater efficiency, (3) an understanding of the interactions between label accuracy and overall system accuracy, and (4) a model for understanding or predicting the expected performance of the overall system. The remainder of this report addresses these key issues.

Section 4 presents research in the area of objective labeling to evaluate automatic and analyst-based technology for attribute assignment.

Section 5 discusses issues related to stratification, sampling, and estimation that considers overall system efficiency.

Section 6 presents both analytic and empirical considerations relevant to developing a generic error model appropriate to SAE technology and specific to Procedure M.

4
OBJECTIVE TECHNIQUES FOR LABELING

The association of ground cover class labels with selected samples of remotely sensed data is an important step in crop inventory applications. Results of the Large Area Crop Inventory Experiment (LACIE) showed that mislabeling of training data was the most important source of error in estimates of wheat and small grain acreages [10]. This should not be too surprising, given the fact that LACIE addressed the problem of estimation in foreign areas and therefore constrained itself to use no direct ground observations of crop type (usually called ground truth) in its labeling process.* This constraint, together with the performance results obtained, points to a need for improved labeling techniques and procedures to minimize errors at this critical step of the area estimation process.

The overall objective of the work reported in this section is to develop a technology to accurately and objectively label scene elements, based on remotely sensed and collateral data. These initial efforts have focussed on wheat and small grains but the general approach should be extendable to other crops as well.

In many respects, the operations of the labeling and classification are similar. Both result in a cover-class or crop label being associated with each element considered, but one can draw a few distinctions. First, labeling is generally more labor intensive than classification. Therefore, labeling is usually performed on a subset of the scene elements whereas classification is more likely to be performed on all or a large fraction of them. Most important, labeling decisions are based primarily, if not completely, on pre-existing decision criteria, whereas classification decisions are based on statistical parameters and measures derived from labeled but localized training samples.

*Ground data were, however, used to assess accuracy of performance and for research and development activities.

Several hypothetical examples can help illustrate these distinctions between labeling and classification. An obvious but clear example is the use of ground observations of crop type to label samples in Landsat data which are then used to train a spectral classifier and classify the remaining elements in the scene as to their crop type. The ground observer's decision criteria are well established before the field visits are made. The labels here happen to be based on information distinctly different and separate from the Landsat data characteristics used to assign the other scene elements to the defined classes. A ground visit to each scene element would be more costly than the classification operations performed on the Landsat data.

As a second example, consider crop labels assigned to samples by an analyst-interpreter using detailed knowledge of the site, multidate Landsat imagery, numerical information extracted from the data, experience in interpreting Landsat imagery and data analysis, and a wide variety of collateral data, such as historical information on local agricultural practices, current-year weather, and knowledge of prior-year crop acreages in the site. These samples can then define a spectral classification rule or be otherwise used to extend the label information to the remainder of the scene. Here, the label assignment is largely based on the Landsat data. However, the analyst-interpreter's interpretation experience and training should have had in place some rather well established decision criteria which could be fine tuned to those conditions indicated by the collateral data and deductions gleaned from the Landsat data. Presumably, the classifier provides a cost-effective and efficient method of extending the label information to the remainder of the scene.

Thirdly, the labeling may be performed primarily by the machine but still require inputs from, and interaction with, an analyst-interpreter who integrates collateral information and makes specific judgements about the overall site characteristics and conditions at computer-specified training locations.

Finally, consider a purely machine labeling operation. The computer accepts the Landsat data and digitized collateral data describing current-year conditions, preferably at a segment or field level. Using this to index a stored bank of data and relationships between historical agronomic data, Landsat data for various crop types, and associated collateral data from prior years and a variety of sites, a probabilistic multiclass label is assigned to each observation. Since the relationships are pre-defined in the data bank and its functional relationships, a labeling decision can be made directly for each scene element, without the need to resort to training and classification, unless dictated by cost considerations.

The above examples illustrate the wide range of labeling situations and requirements that may exist or evolve as technology develops.

4.1 BACKGROUND

In LACIE, labeling within 5x6-mile segments was performed by analyst-interpreters using false-color images from multiple acquisitions of Landsat data throughout the growing season as their primary source of identification information. Collateral data in the form of topographic maps, weather, and historical agricultural statistics (e.g., at a county or crop-reporting-district level), and historical crop calendars were also available. Throughout the three phases of LACIE, the amount and quality of collateral data improved. In Phase III, displays of numerical information extracted from Landsat digital data found limited use by analyst-interpreters and selected full-frame Landsat images were available to provide a broad geographical perspective for interpreting the segment data.

During Phases I and II, analyst-interpreters were required to select, delineate, and then label a sample of fields that was supposed to span the range of spectral variability in each segment. For Phase III, Procedure 1 was developed and used. With Procedure 1, analysts label "dots", that is, individual pixels located at intersections of a ten-line-by-ten-column grid superimposed on the segment. Two samples of these dots are labeled, Type 1 dots and Type 2 dots. All labeled Type 1 dots lie away from field edges and boundaries, in the analyst-interpreters judgement, while Type 2 dots can fall on edges or boundaries as well as in field centers. Type 1 dots are used to label clusters while Type 2 dots are used for bias correction, i.e., to form a stratified area estimate as described earlier in Section 3.

Accuracy assessment analysis of LACIE Phase III results [10] identified four sources of labeling error. The first was "abnormal signatures" which refers to fields that did not follow the temporal sequence expected for their crop type under conditions believed by the analyst to be occurring in the segment. Boundary and edge pixels were another major source of error for Type 2 dots; although more common in strip-farming areas, they can be plentiful in other areas as well. A third source of error was inadequate acquisition history; when one or more key acquisitions were missing, large errors did result on occasion. Other types of errors included clerical errors and inconsistent labeling of pixels having the same temporal sequence in the same segment. One particularly difficult discrimination problem for analysts was distinguishing spring wheat from barley and other spring small grains.

Image interpretation is a complex process that depends on several factors: (1) the interpreter's training, background, and experience, (2) the extent to which collateral and supporting image data place the interpreter in the proper context, (3) the number of characteristics or features that are deducible from the available data for each scene element to be identified and/or described, and (4) the detection or

definition of those labeling targets. These factors, as well as the labeling difficulties found in LACIE, must be addressed in the development of improved procedures and techniques.

Several research and development activities during LACIE and subsequently during LACIE Transition have been addressing many of the above labeling problems. At ERIM, as reported last year [3, 11], three aspects were addressed. First, an overall area-estimation system framework was developed (see Section 5). Called Procedure M, it includes a sequence of modules to condition, stratify, and sample the data and produce estimates. From the labeling standpoint, its major differences from Procedure 1 are its preprocessing operations which stabilize and transform the data and its computer definition of quasi-field shapes for labeling. The second aspect addressed was that of improved color products for use in labeling by analyst-interpreters. The third aspect addressed at ERIM was a machine labeler for discriminating between spring wheat and barley. It features a crop-calendar shift calculation based on the temporal profile of a vegetation-indicating variable derived from Landsat, i.e., the Greenness variable of the Tasseled-Cap Transformation [12, 13]. This labeler was tested on ground-truth-labeled (small grain vs. non-small grain) quasi-fields. While performing well on conditions similar to those used in its development, needs for refinement to handle other conditions and for testing with analyst-interpreter grain vs. non-grain labels were identified.

At JSC, development of the LIST (Label Identification from Statistical Tabulation) procedure was carried out [14]. In this procedure, the machine makes the labeling decision utilizing analyst inputs together with machine-derived variables. The key AI questions and machine variables were determined through stepwise regression analyses. The questions answered by the analyst-interpreter about each selected dot are: (a) Is the pixel clearly in a non-agricultural area? (b) Is

the pixel registered in all Landsat acquisitions used? (c) Is the pixel a mixed pixel? (d) Is this an anomalous pixel (with respect to others in the field)? (e) What indication of vegetation canopy density and condition is present on the imagery for each date (one of six categories)? The machine computes a Green Number* and a Brightness for each date as well as multitemporal greenness image eigenvector components [14], compares them to expected patterns as a function of Robertson biostage for wheat, and makes its decision using both the analyst-interpreter and the machine inputs.

Two other related studies were pursued at JSC. First, as a segment-level spectral indicator of drought stress, the Green Index Number (GIN) was developed and successfully tested in the U.S. Great Plains [16]. Second, the importance of crop calendar differences and time profiles was identified and pursued in a machine classification context [15]; performance comparable to analyst labeling was achieved in initial tests.

At the University of California at Berkeley (UCB), a "delta function" stratification procedure was developed to assist analysts [17]. This procedure stratifies unsupervised Landsat spectral clusters according to the temporal pattern observed in a quantized form of the ratio of two Landsat bands, MSS7/MSS5, which is highly correlated to vegetative development. The quantization is primarily at the level of vegetation vs. no vegetation on each date. Estimates of vegetation cover density also are utilized by analysts.

*Note that "Green Number" [15] is similar to but not identical to "Greenness" which results from application of the Tasseled-Cap transformation [12].

4.2 APPROACH

As illustrated by the examples at the beginning of Section 4, labeling processes could run the gamut from a purely manual function to a purely machine (i.e., computer) function. In LACIE, labeling was primarily manual with some machine assistance. Purely machine techniques appear to be at least several years away, so our intent is to concentrate on optimizing the balance between man and machine, with the machine being asked to do as much as it can do acceptably well. Another key focus is making maximum use of the multidate aspects of Landsat and collateral data.

The approach being taken is an iterative process involving four types of activities. The first is feature definition based on relationships between observable spectral characteristics of crops and factors that affect these observables. An integral part of feature definition, then, is analysis of crop physiological, agronomic, and physical characteristics and relevant collateral data to obtain a better understanding of those relationships. The second type of activity is feature extraction. Features are derived and extracted manually or by machine from spectral and collateral data and, ideally, enhance or simplify the identification and discrimination of crops. The third activity is characterizing, for each given crop, the statistical properties and physical relationships of features and interactions between them. This we call signature characterization. Finally, there is procedure development, that is, the utilization of crop signature characterizations to define (preferably objective) rules for assigning crop labels based on observed features. These rules must be incorporated into labeling procedures that can be implemented, tested, and evaluated for use in area estimation systems.

During this first year, efforts have been directed mainly at greater understanding of features, analyst-interpreter functions, and

machine-related functions and at developing tools for increasing this understanding.

4.3 GENERAL DISCUSSION OF FEATURES

Features are data properties or characteristics that help to detect, assess the condition of, and distinguish between the various crop types and cover classes present in remotely sensed data. When two crops have sufficiently different values for a given feature or set of features, they can be distinguished from one another. However, statistical discriminability is not sufficient. To be truly useful in a labeling context, features must also be relateable in some predictable manner to inherent physical, physiological, or agronomic characteristics of classes of interest. "Features" as used in this report, imply such a relationship.

Remotely sensed data provide three basic dimensions for feature building--spatial, spectral, and temporal. Features may relate to individual pixels, individual fields, collections of pixels or fields, or entire segments. In addition, collateral data may provide additional information which allows for conditional interpretation of a particular feature to assist in discrimination and identification.

Feature building from remote sensing data and collateral data is discussed in Section 4.3.1. Section 4.3.2 then discusses some of the relationships of features to agronomic and physiological characteristics. Next, methods of extracting features from available data and potential uses are discussed in Section 4.3.3.

4.3.1 TYPES OF FEATURES

The field pattern in agricultural areas is the most obvious spatial feature which can be used to detect and determine field boundaries, fields being very important entities for agricultural inventory and

and condition assessment. The field pattern includes size, shape, and texture information, as well as a spatial context for interpretation. As noted earlier, LACIE analyst-interpreters delineated and labeled fields on imagery during Phases 1 and 2, but only had to label individual dots in Phase 2. Procedure M, on the other hand, incorporates a machine operation to delineate quasi-fields for labeling.

Spectral features for individual pixels or fields are the most important features for crop identification and begin with the spectral channels of the sensor. In LACIE, the key information for analyst-interpreter labeling was carried in the colors of film imagery, i.e., representing information in three of the four Landsat bands, and in numerical and graphical displays. In computer classification, all available spectral channels are commonly used to extend training labels to other parts of the scene. The Tasseled-Cap Transformation, on the other hand, defines new features which are linear combinations of the Landsat channels and offer advantages in dimensionality reduction and physical interpretability. Another type of spectral feature is a segment-level indicator such as the previously referenced Green Index Number indicator of drought stress.

We note that, whereas the basic crop information in Landsat spectral data derives from relatively stable scene reflectance properties, the actual observed data are susceptible to changes caused by environmental and observational effects and properties of the display medium. For example, atmospheric conditions and observation angles may differ from location to location and day to day, and image colors may depend on scene content. Normalizing transformations or features that are insensitive to such variations are therefore desirable.

The 18-day (nine-day with two satellites) coverage cycle of Landsat opens up the temporal dimension for feature definition. Crop phenology is a key characteristic which is and can be used to improve identification and discrimination. With multiple acquisition dates,

one can establish or fit temporal-spectral profiles or trajectories to the data and extract new features. One useful feature is the time shift of the profile relative to a standard profile--this gives a measure of local crop calendar shift due to variations in planting date or other factors. A variety of other features derivable from temporal-spectral profiles are discussed in detail in Section 4.4 and Appendix A.

In addition to features based purely on remote sensing inputs, collateral data may provide feature-conditioning information useful in identification and discrimination. General weather patterns can affect all fields in a segment, as in a drought situation. Soil and terrain characteristics can and do affect individual fields differently. While not attainable now, meteorological information at the segment or even approaching the field level may become available from meteorological satellites in the future.

4.3.2 FEATURES AND AGRONOMIC CHARACTERISTICS

Basic characteristics of crops that vary throughout the growing season are the type and amount of vegetation covering the soil and the condition, vigor, and greenness of that vegetation. Landsat responds to the overall combination of soil, vegetation, and plant geometry, as influenced by solar and viewing geometries.

A number of features have been developed and used by different investigators to indicate the amount of green biomass, green leaf area, or green ground cover present. Appendix B discusses several. We have chosen to use the Tasseled-Cap Greenness variable as our primary indicator of growing vegetation. It is largely insensitive to soil color effects and is a linear combination of the original Landsat bands. For small grain crops, it reaches a peak value just prior to heading.

Since vegetation indicators or derivatives thereof also respond to crop condition, they may be used to monitor crop condition and provide early warning of anomalous conditions. At the segment level, we have discussed the GIN drought stress indicator. In Section 4.8 and Appendix G, a more specific form of segment-level stress indicator is developed for small grains, as a refinement of our machine labeler for spring wheat. This indicator is based on fitting a temporal-spectral profile to small grain training data.

While most vegetation-indicating features are relatively insensitive to soil brightness, one can define other features that respond well to changes in soil brightness (given a low percentage of covering vegetation). Tasseled-Cap Brightness is one such example. As discussed in Section 4.8, use of temporal-spectral profile fitting provides one way, an automatic one, of determining which acquisitions of small grains correspond to bare soil. A second refinement to the spring wheat labeler alters the decision rule depending on the soil brightness in the segment.

4.3.3 EXTRACTION AND USE OF FEATURES

Features may be extracted by either an analyst-interpreter or a machine or by an interactive combination of the two. This extraction may be implicit as well as explicit. For example, an analyst-interpreter may implicitly use field features when labeling individual pixels or dots. Many of the features extracted by analyst-interpreters are amenable to machine extraction as well.

Features have several possible uses. The first is as a direct basis for labeling. Second, they may be aids or collateral inputs to analyst-interpreters in their labeling and interpretation.

Third, they may be variables used in stratification and/or classification. Fourth, they may be major sources of information for early warning and assessment of anomalous crop conditions. Finally, they may serve as spectral inputs to crop yield estimation models.

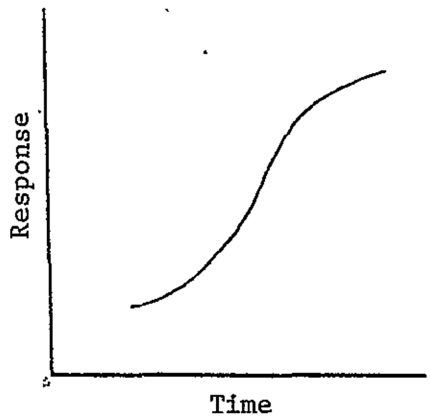
4.4 FEATURES BASED ON TEMPORAL-SPECTRAL PROFILES

Because of their potential importance and the amount of attention they received from us during the course of the year, this section elaborates on the role of temporal-spectral profiles or trajectories as sources of features. By "profile" we mean a mathematical representation of the temporal pattern of a given variable or variables for a given crop.

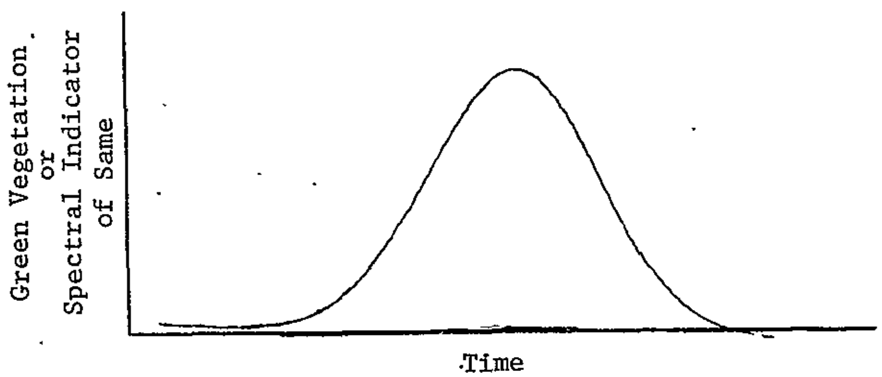
4.4.1 GENERAL DISCUSSION

The underlying assumptions of profile modeling and fitting are that (1) crop development and accompanying spectral changes are continuous processes with gradual rather than abrupt changes, (2) both processes can be characterized by a combination of Sigmoid-shaped curves which are typical of most biological phenomena, and (3) characteristic spectral development patterns exist for specific crops or groups of crops. Illustrations of a Sigmoid curve and green vegetation development profile are presented in Figure 3.

Next, consider a specific green vegetation indicator, namely a Greenness variable. Part (a) of Figure 4 illustrates a smoothed time profile for wheat, as extracted from field-measured reflectance data (see Appendix B). The corresponding profile for a Brightness variable is shown in Part (b). Between Tasseled-Cap Brightness and Greenness, the vast majority of Landsat data variability on any one date is captured. By combining the information in Parts (a) and (b), Part (c) illustrates the spectral path or spectral track followed by the test field of wheat during the growing season. Time is the third dimension

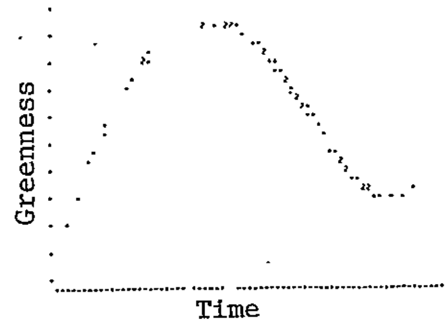


(a) Sigmoid Curve Typical of
Biological Processes

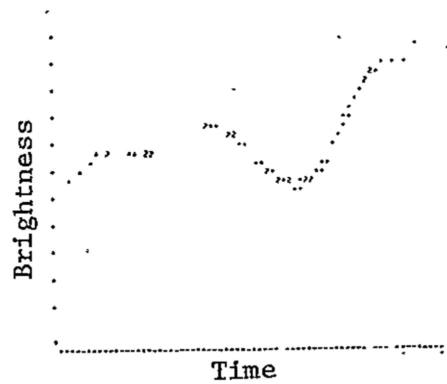


(b) Crop Development Profile

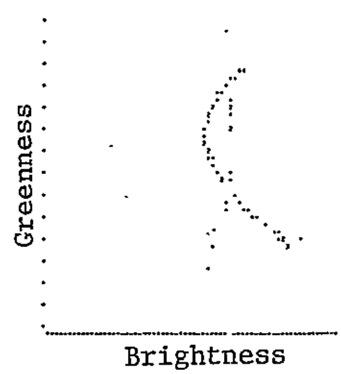
FIGURE 3. ILLUSTRATIONS OF CURVE SHAPES



(a) Greenness Profile



(b) Brightness Profile



(c) Spectral Path or Track

FIGURE 4. ILLUSTRATION OF TEMPORAL-SPECTRAL PROFILES IN REFLECTANCE SPACE FOR A TEST FIELD OF WHEAT

along which one can imagine a temporal-spectral trajectory followed by any given field or crop.

For this report, we limit consideration to spring small grain crops and to spectral indicators of green vegetation. However, the approach should be adaptable to other crops with a single green-up phase and to other spectral variables. Detailed discussion of profile modeling and fitting is presented in Appendix A. In Appendix B, a variety of green vegetation measures are discussed. However, for the remainder of this section, we use Tasseled-Cap Greenness as the profile variable.

One implication of profile modeling and fitting is that the actual continuous crop development patterns can be more accurately characterized from a set of intermittent observations made by Landsat or other remote sensors. Another is that more detailed estimates may be possible of crop development stages and of the influence of crop stresses and other factors on development patterns.

4.4.2 LEVELS OF USE

Some of the many possible applications of temporal-spectral profiles have already been mentioned. These different applications do, in general, put different demands on model forms used to characterize the profiles. To help focus on these differences, four levels of use have been defined. These levels and their requirements are summarized in Table 5 and discussed in Appendix A. Basically, they require progressively more exact characterization of the profile shape. The first level, which has received most use, computes a crop calendar shift by sliding either the given set of multidate observations or the reference profile past the other in time until they best match.

TABLE 5. LEVELS OF USE OF TEMPORAL-SPECTRAL PROFILES

Level	Requirement	Example Uses
1 - General Form	Rough approximation of temporal development, e.g., linear interpolation or functional fit	Crop Calendar Shift Estimation Stratification Variable
2 - Estimation of a Particular Spectral Feature	Accurate representation of profile portion related to the feature, not necessarily whole profile	Estimation of Peak Greenness Estimation of specific crop development stage
3 - Characterization of Multiple Crop Features, Overall Spectral Development	Accurate fit to data throughout growing season	Extraction of new or added features for labeling, crop condition assessment, etc.
4 - Substitution of Missing Data	Very accurate fit; mechanism for sensible interpolation based on collateral data and understanding	Characterization of a general signature for wheat

4.4.3 MODEL FORMS

Requirements on model forms and fitting procedures differ with the level of application, as noted in Table 5. The least stringent requirement is placed on Level 1, exemplified by the crop calendar shift calculations inherent in our machine labeling procedure described in Section 4.8 and Appendix G. In a pioneering work on crop calendar shift calculation [16], straight line segments were used to connect means for the acquisition dates of selected training fields. We at ERIM introduced the use of a model to provide a better representation of profiles, to help compensate for missing acquisitions, and, with a logarithmic transformation, permit a multiple linear regression calculation of parameters [18, 3]. The model form was:

$$F(t) = at^{b}e^{ct^{2}} \tag{1}$$

where $t = t'-t_o$ = shifted days from reference date,

t' = day of year after crop calendar shift,

t_o = reference day,

$F(t) = G(t)-G_o$ = Greenness value above offset,

$G(t)$ = Greenness value,

G_o = Greenness offset,

a,b,c = model parameters.

This function has a peak value

$$F(t_p) = a\left(\frac{b}{-2ce}\right)^{b/2}$$

which occurs at

$$t_p = \sqrt{\frac{b}{-2c}}$$

This model form has served well for Level 1 applications. Additional discussion of its use and fine points of its application are presented in Section 4.6 and Appendix A; a key point is the importance of using the offsets t_o and G_o.

In moving to Level 2 and Level 3 applications, it was found desirable to develop a new, but related, model form that reduced the importance of offsetting and gave better and independent fitting of the data values before and after the peak, while still ensuring continuity at the peak. This new model form, applied after first computing the day of peak Greenness, is:

$$F(t) = \begin{cases} ae^{b_1(t-t_p)^2} & ; \; t > t_p \\ ae^{b_2(t-t_p)^2} & ; \; t \leq t_p \end{cases} \tag{2}$$

where $F(t)$ = Greenness-G_o,
t = day of acquisition,
t_p = estimated day of peak Greenness,
a, b_1, b_2 = parameters to be estimated.

This function has a peak value of "a". Advantages and applications of this model form are discussed in Section 4.6 and Appendix A.

4.5 STUDIES FOR INCREASED UNDERSTANDING OF TEMPORAL-SPECTRAL CHARACTERISTICS

Both empirical and theoretical approaches are desirable for understanding the temporal-spectral characteristics of crops. Empirical studies make use of data bases that assemble and correlate Landsat and field measurement data with agronomic and other collateral variables. Optimal utilization of such data bases depends on understanding the structure of the data and features in spectral space; Tasseled-Cap Transformation features are an integral part of most of our research.

Initial results of a study of spectral space relationships are presented in Section 4.5.1. In addition, Section 4.5.2 reports on an analytic modeling effort that was initiated. The objective of that effort is to develop, as an analysis tool, a capability to model the spectral phenology of wheat, incorporating meteorological inputs and predicting Landsat variables.

4.5.1 LANDSAT SPACE, REFLECTANCE SPACE, AND THE TASSELED-CAP

The objectives of this study are two-fold. The first is to better understand the relationships between data values and their four-dimensional structure in Landsat spectral space and in corresponding reflectance space. The second is to acquire insights about the information content of Landsat data pertaining to crop development and identification. In the future, similar consideration of Thematic Mapper relationships is desired. Progress and results achieved to date are discussed separately for Landsat and reflectance data.

4.5.1.1 Landsat Data Relationships

For analysis and processing, we recommend and routinely do make adjustments to Landsat data values to correct for difference in atmospheric haze content, sun angle, and Landsat sensor calibration, using procedures described in Section 5.2. This type of normalization stabilizes the spectral variables and permits a more meaningful and consistent use and interpretation of multidate Landsat data. The relationships given in this section should not be applied directly to data that have not been similarly corrected, but they should still be useful in general interpretation of such Landsat data.

The plane formed by the Brightness and Greenness variables of the Tasseled-Cap Transformation contains the vast majority of Landsat data variability and facilitates physical interpretation of data values and spectral trajectories. However, relationships between this plane and the original Landsat variables and features, such as the ratios MSS5/MSS7 and MSS4/MSS5, are not easily visualized from the transformation equations.

To assist this visualization, three graphs have been prepared. For reference, each contains an outline of the typical distribution of data from agricultural scenes as well as some threshold lines used to screen the data prior to correction.

Figure 5 shows a line for each Landsat band. Each line represents the locus of all points in the plane when that band value is zero. The correlations between Bands 4 and 5 and between Bands 6 and 7 are evident in the similarity of their slopes. Also shown by each line is a parallel dashed line that represents the effect that a ten-count increase in the band value would have on the line's position. Thus, one could conceivably build up grids to overlay on the Tasseled-Cap plane and read off Brightness and Greenness values for any point in XSTAR-corrected Landsat space.

Note that the lines for the four bands do not intersect at a single point. The reason is that the Tasseled-Cap plane does not pass through the origin of Landsat space. It is displaced by -11.21 counts in Yellow and 1.356 counts in None-such, the third and fourth components of the transformation.

Lines for fixed values of the Band 7 to Band 5 ratio are presented in Figure 6. This ratio is commonly used as a vegetation-indicating feature. The line shown for R=0.55 corresponds to a particular threshold,

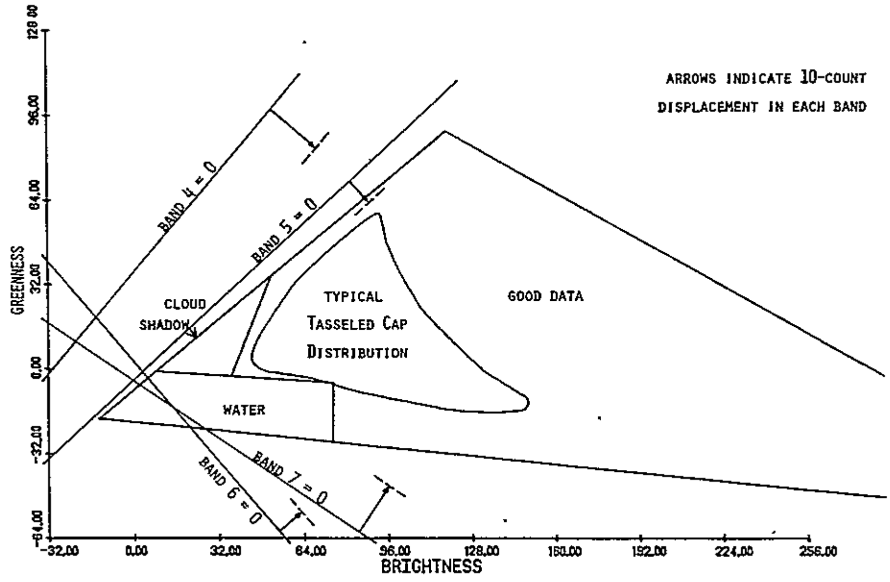


FIGURE 5. LANDSAT BAND VALUES REFERRED TO XSTAR-STANDARDIZED TASSELED-CAP PLANE

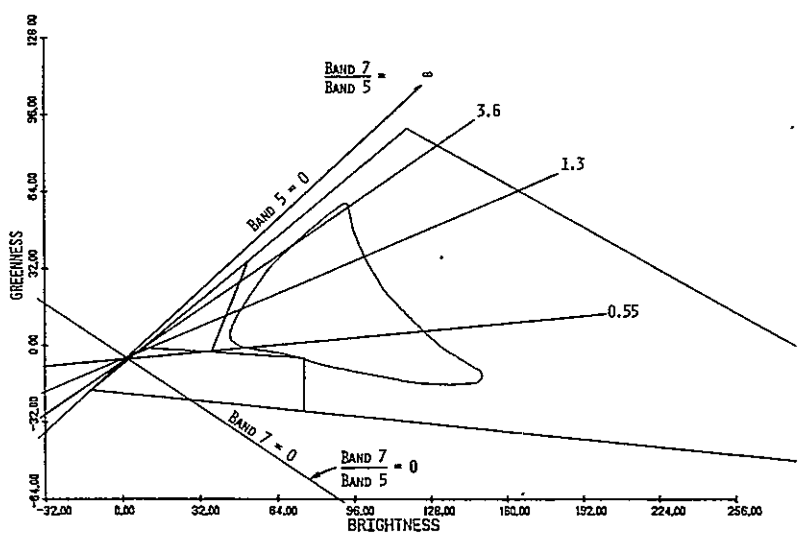


FIGURE 6. LANDSAT BAND 7 TO BAND 5 RATIOS REFERRED TO THE XSTAR-STANDARDIZED TASSELED-CAP PLANE

2 x (Band 7/Band 5) = 1.1, that was empirically established and is used at the University of California at Berkeley to indicate presence or absence of green vegetation [17]. The non-linearity of the ratio variable is evident from the spacing of the lines and their values.

Figure 7 presents lines for fixed values of the Band 4 to Band 5 ratio, a ratio that has received some attention as a possible indicator of crop development stage. Lines for different values of this ratio do progress across the Tasseled-Cap plane largely parallel to those of the individual Bands 4 and 5. The distance-from-green-arm measure used in our Spring Wheat vs. Barley labeler also performs a similar partitioning, with the advantage of somewhat reduced noise effects due to band averaging. This ratio, however, does not appear to be a good indicator of green development above the predominantly horizontal line or band of bare-soil signal values. The importance of atmospheric and other corrections to stabilize the data is reemphasized because of the large atmospheric contribution to Band 4 values, particularly, and to Band 5.

Experience has shown the Yellow component of the Tasseled-Cap transformation to be primarily sensitive to changes in atmospheric haze. Therefore, the sensitivity of Brightness and Greenness to changes in Yellow for fixed values of the two ratios were determined analytically (see Part D of Table 6). In the upper-left half of the Tasseled-Cap distribution of agricultural data, the sensitivity to changes in Yellow computed for the MSS4/MSS5 ratio is three times greater than that for the MSS7/MSS5 ratio and is roughly three counts per count of change in Yellow. Thus, we see that the MSS4/MSS5 ratio is very sensitive to external non-agronomic effects.

Mathematical expressions for the lines displayed in Figures 5, 6, and 7 are presented in Table 6; the Tasseled-Cap Transformation is repeated in Table 7 for completeness [19]. Again, we caution that they refer to XSTAR-corrected data.

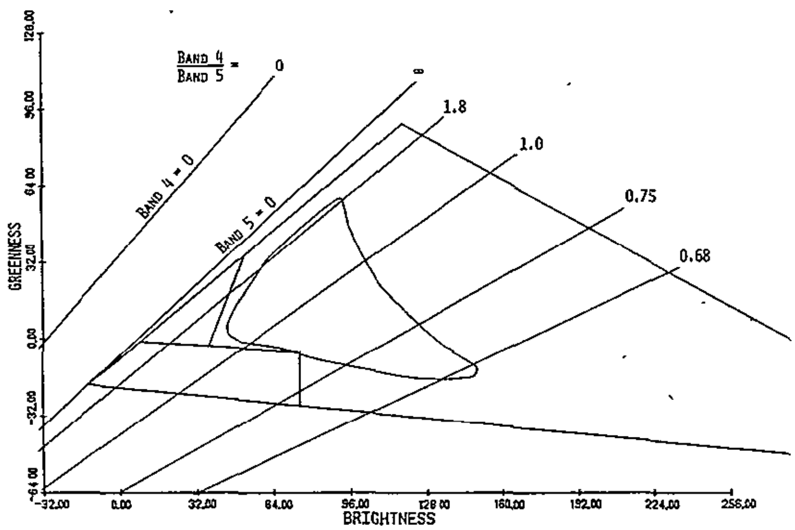


FIGURE 7. LANDSAT BAND 4 TO BAND 5 RATIOS
REFERRED TO THE XSTAR-STANDARDIZED
TASSELED-CAP PLANE

TABLE 6. MATHEMATICAL EXPRESSIONS FOR LINES IN THE XSTAR-STABILIZED TASSELED-CAP PLANE

A. LINES FOR INDIVIDUAL LANDSAT BANDS*

$$G = 1.1735\,B + 35.5277 - 3.53145 * MSS4$$
$$G = 0.9138\,B - 7.0044 - 1.5150 * MSS5$$
$$G = -1.1705\,B + 3.1310 + 1.7321 * MSS6$$
$$G = -0.6767\,B - 4.2581 + 2.5751 * MSS7$$

B. LINES FOR RATIOS OF LANDSAT BANDS

$$x = \frac{MSS7}{MSS5} : \quad G = \left[\frac{0.60316(x) - 0.26278}{0.66006(x) + 0.38833}\right] B - \left[\frac{4.6233\ (x) + 1.6535}{0.66006(x) + 0.38833}\right]$$

$$y = \frac{MSS4}{MSS5} : \quad G = \left[\frac{0.60316(y) - 0.33231}{0.66006(y) - 0.28317}\right] B - \left[\frac{4.6233\ (y) + 10.0604}{0.66006(y) - 0.28317}\right]$$

C. LINES FOR RATIO PASSING THROUGH A SPECIFIED POINT IN PLANE

$$\left(\frac{MSS7}{MSS5}\right)_o = \left[\frac{0.26278\ B_o + 0.38833\ G_o + 1.6535}{0.60316\ B_o - 0.66006\ G_o - 4.6233}\right]$$

$$\left(\frac{MSS4}{MSS5}\right)_o = \left[\frac{0.33231\ B_o - 0.28317\ G_o + 10.0604}{0.60316\ B_o - 0.66006\ G_o - 4.6233}\right]$$

D. SENSITIVITY OF RATIO VALUES TO MOVEMENT IN YELLOW DIRECTION

$$x = \frac{MSS7}{MSS5} : \quad \frac{dG}{dY} = \frac{0.43\,x - 0.04}{0.66\,x + 0.39} \quad \text{and} \quad \frac{dB}{dY} = \frac{0.43\,x + 0.04}{-0.60\,x + 0.26}$$

$$y = \frac{MSS4}{MSS5} : \quad \frac{dG}{dY} = \frac{0.43\,y + 0.90}{0.66\,y - 0.28} \quad \text{and} \quad \frac{dB}{dY} = \frac{0.43\,y + 0.90}{-0.60\,y + 0.33}$$

*Note: G = Greenness and B = Brightness, as defined in Table 7.

TABLE 7. TASSELED-CAP TRANSFORMATION COEFFICIENTS FOR LACIE-CALIBRATED LANDSAT 2 DATA [20]

$$\begin{bmatrix} \text{Brightness} \\ \text{Greenness} \\ \text{Yellow} \\ \text{Non-Such} \end{bmatrix} = \begin{bmatrix} 0.33231 & 0.60316 & 0.67581 & 0.26278 \\ -.28317 & -.66006 & .57735 & .38833 \\ -.89952 & .42830 & .07592 & -.04080 \\ -.01594 & .13068 & -.45187 & .88232 \end{bmatrix} \begin{bmatrix} \text{MSS4} \\ \text{MSS5} \\ \text{MSS6} \\ \text{MSS7} \end{bmatrix}$$

Note: To avoid negative values in data processing, it is common practice to add 32 counts to each of the Tasseled-Cap variables. This was done to the Greenness variable used for profile matching in this report (Sections 4.4 and 4.8 and Appendices A and G).

4.5.1.2 Reflectance Data Relationships

Field measurements of reflectance either within or integrated over Landsat bands are a valuable source of information and understanding. Appendix B describes methods we have used to analyze this type of data and to establish a Tasseled-Cap-like transformation for them. Both principal component analysis and approximate Landsat band-to-band calibration ratios were employed.

Analyses did show that a very high percentage (98 to 99%) of data from a full-season measurement series on wheat plots was in a two-dimensional plane. Senescing vegetation was not found to lie off that plane by an appreciable amount.

4.5.2 MODELING OF CROP SPECTRAL PHENOLOGY (SEED-TO-SATELLITE MODEL)

Crop labeling procedures can be better developed and field measurements and Landsat data acquired intermittently over a growing season can be better analyzed with a good understanding of the biophysical processes that underlie the crop spectral phenomena. Furthermore, there is a need to be able to explore the effects of varied parameters on crop development under a greater variety of meteorological and other conditions than those for which detailed measurements exist. Therefore, a long-term effort was initiated to develop a capability to analytically model the spectral phenology of wheat.

The planned approach is to first formulate an overall structure for the model and its components and then acquire state-of-the-art (SOA) or near-SOA submodels for those components. The components will be implemented and tested individually, followed by development of the required interfaces. Input and test data sets will be acquired and the full model will be synthesized and tested. Updates will occur as improved submodels become available. The overall model has been nicknamed the "Seed-to-Satellite" model.

During this reporting period, the overall structure shown in Figure 8 was established. It begins with a meteorologically driven wheat growth model which interfaces with a crop reflectance model and then the remote sensing submodels to represent atmospheric and solar effects, the sensor, and preprocessing and feature extraction.

Contact was made with Dr. J. T. Richie of the USDA and versions of his wheat growth model [20] were obtained and made operational on our computer. We plan to develop interfaces between it and the bi-directional crop canopy spectral reflectance model which was developed by G. Suits at ERIM [21]. This in turn will be linked to the radiative transfer model developed by R. Turner while at ERIM [22].

We also hope to acquire two to three decades of meteorological data over agricultural regions and produce a principal component characterization of that data. Such a characterization should provide a baseline weather scenario and a limited number of deviant scenarios that are representative of the major variability to be encountered.

4.6 IMPROVEMENTS IN FEATURE EXTRACTION

The extraction of features from Landsat and collateral data is important to crop identification and inventory applications. Five steps supporting improved feature extraction were made during the year, the first two being primarily oriented toward aiding analyst-interpreters in their labeling and the others having potential application in both manual and machine approaches to labeling. The first two topics discussed below are development of aids for the labeling of blobs (quasi-fields) and the development of color-stabilized image products. The other three topics are improved spatial-spectral clustering, estimation of crop development stage using Landsat, and investigation of temporal-spectral profile fitting procedures.

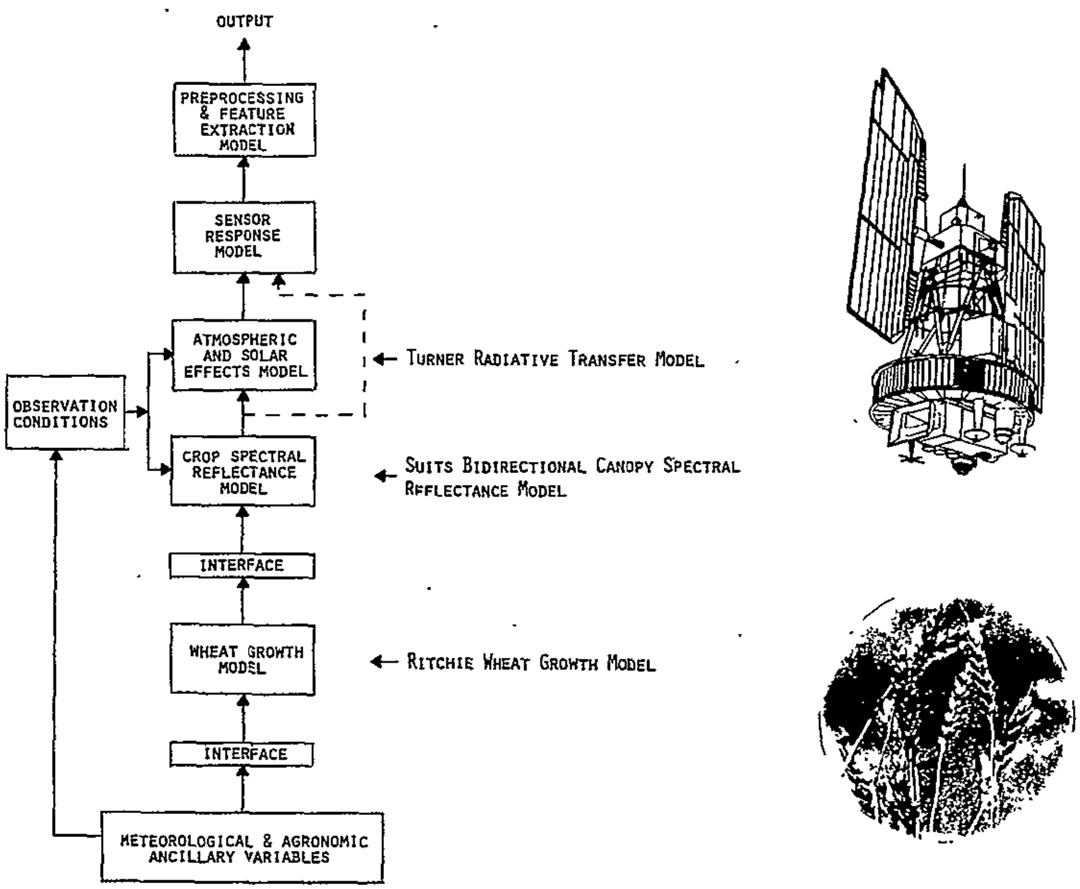


FIGURE 8. OVERALL STRUCTURE OF THE "SEED-TO-SATELLITE" MODEL FOR WHEAT

4.6.1 DEVELOPMENT OF AIDS FOR LABELING BLOB TARGETS

The analyst-interpreter labeling experiment discussed in Section 4.9 represents the first experience that has been gained with manual labeling of the field-like blobs or quasi-fields defined by our spatial-spectral clustering algorithm, BLOB. Some rudimentary labeling aids had to be produced in order for the experiment to be conducted. Time permitted only a minimal development; therefore, substantial additions and improvements would be highly desirable in the future. For instance, no spectral aids were produced.

The BLOB algorithm assigns each pixel to a blob and a unique number to each blob. The major problems addressed by the labeling aids were those of transferring information about this blob pattern to the analyst-interpreter for use in labeling and providing a means for the labels to be recorded and digitized. Another operation performed by the BLOB algorithm is a mathematical stripping of the edge pixels from the periphery of each blob, permitting the designation of blob centers. Only blobs large enough to have interior pixels were to be designated to analyst-interpreters for labeling.

Two types of image products were developed. The first was a black and white film mask to delineate blob interiors when placed over standard LACIE segment images and printed by the same system (the JSC Production Film Converter) as the LACIE segment images. The second type of image product was a corresponding digital line-printer map that served both as a medium for recording label assignments and for providing additional detail about blob number assignments.

Four blob interior masks were produced for each segment for each analyst, to permit simultaneous use on images acquired on different dates or on different image products. All but blob interiors were made

opaque. The line-printer map had blanks for blob interiors, an over-printed symbol encoding the blob number for edge pixels around each interior, and another symbol (the period) for pixels in blobs that did not have interiors. A color-coding scheme was devised by the analyst-interpreters for recording labels, as discussed in Section 4.9.

Later, for use in digitizing the color-coded labels, a computer program was written to produce a listing of blob numbers, ordered by line/pixel coordinates of the first interior pixel encountered. We believe this would have been a more convenient medium for analyst-interpreters to record their labels and its use would reduce the chances of skipping or overlooking small blobs.

As noted earlier, time constraints precluded the development of additional labeling aids such as the spectral plots produced for dots in LACIE Procedure 1. Such plots and other more advanced aids could be produced in a straightforward manner for blobs and are recommended for future experiments. We also believe that development of a capability to outline quasi-field boundaries on the image overlays would be a marked improvement and prove less of an obscuration of scene detail than did the masks used in this experiment.

4.6.2 COLOR-STABILIZED IMAGERY

Imagery produced by the Production Film Converter (PFC) has played a crucial role throughout LACIE in the estimation of crop acreage. Imagery is the basic product utilized by analyst interpreters in labeling samples. The techniques used in LACIE for producing color imagery from Landsat data were designed to imitate imagery from color infrared aerial photography, as do the conventional Landsat image products. This filled the need for imagery understandable by conventional photo-interpretation. In essence, each of three Landsat data channels was

used to control the exposure of film in a different primary color. The radiometric image of the scene as viewed by Landsat was written on the image three times, each time using a different band and each time using a different primary color. The results were imitations of what one would obtain directly by photographing the scene using color infrared film with appropriate filters and spectral sensitivities.

The majority of labeling in LACIE was performed using LACIE Product 1 images [23], with supplemental use of Product 3 images [24]. Other image production algorithms also have been considered [23, 25, 26]. These algorithms form a family of approaches that are characterized principally by the following:

(a) A subset of the four MSS channels are projected into PFC gun counts.

(b) Each channel undergoes mean adjustment and contrast stretching. Each algorithm considers the dynamic ranges of the MSS channels and manipulates them by scaling and biasing (affine transformations) according to some rule, generally to fill out the range of color intensities the photographic system can produce.

Image products based on the 'bias and scale' approach have served and continue to serve needs for image interpretation. Yet it has been recognized that they are not without shortcomings. The Product 1 was found to produce color variations from segment to segment and from date to date that were not consistent in meaning. Product 3 addressed this in part by providing an improved level of consistency in hue and saturation, but often at the cost of adequate color resolution.

ERIM experimentally quantified [27] shortcomings of these bias-and-scale approaches, illustrating that:

(1) Significant information is lost in omitting a spectral channel.

(2) Bias and scale approaches cannot provide consistent color meaning for they cannot truly adapt to the natural variations in atmosphere and sun position that cause color inconsistencies, and they falsely adapt to variations in scene composition.

(3) Perceptual considerations in the interpretation of colors mapped by gun count projection can result in a distortion of the actual relationships among data values.

4.6.2.1 General Approach

In this section, an approach to image product production is presented that is based on an entirely different principle than the products discussed above. The new principle is that Landsat data features should be projected into color space directly, rather than into PFC gun counts. This concept was utilized by M. Walker [28] in 1974 in producing experimental imagery by a technique which mapped the principle components of Landsat data into an 'opponent poles' color space. R. Juday [29] of JSC/EOD in 1977 began work that was based on a similar approach which utilized for the first time a model of the photographic system (PFC) and a perceptually uniform color space in the mapping.

Research at ERIM, motivated by the work of Juday, introduced other considerations based on an understanding of the sensor system and the structure of Landsat data. The underlying principle rests in mapping surface reflectance values (as opposed to radiance values) in a consistent manner to well defined colors.

This ambitious objective is addressable by the following threefold approach:

(1) Take advantage of our knowledge about Landsat data to use imagery to exploit the information contained therein. Derived variables representing features of interest in the data are used to drive color variation. This is in contrast to the current technique of choosing three out of the four Landsat bands for generating an image. In actuality, two linear combinations of Landsat variables, Greenness and Brightness, span nearly all the variation in Landsat data for an agricultural scene. These features have an interpretation in terms of plant cover and soil background [13]. The Brightness and Greenness features have proven exceedingly useful for simplifying computer classification of agricultural scenes. Adoption of these variables as the driver of color variation promises to simplify image interpretation as well, and yet retain as much or more of the information relevant to interpretation as current products. The third linear combination, Yellow, represents most of the remaining variability and is sensitive to changes in atmospheric haze content and is used as a diagnostic feature.

(2) Utilize data normalization technology to provide a basis for consistent color mapping. A standardized projection of Landsat data that is independent of scene content and adapts for atmospheric conditions and solar geometry provides the framework for mapping to a standard color domain.

(3) Take advantage of Color Science findings to present the imagery in a manner that does not visually distort true relationships among data and at the same time provides adequate color resolution and tone for interpretation. Two essential concepts are involved. First, we know color perception recognizes three components of color or three independent dimensions of color variation. They are hue, lightness,

and saturation. We match our data features to these dimensions of color variation. In this way, the independent data features remain independent to the perception of the image analyst. Second, the judgement of differences between color is amenable to mathematical modeling. Color scientists conceive of a "Uniform Color Space" in which colors are arranged so that Euclidean distance between them is a consistent measure of the difference between them as perceived by the human eye. In order to make imagery we can conceive of mapping our data space linearly into Uniform Color Space. We then expect the quantitative distance relationships of data points to be translated to appropriate color differences, providing to the analyst an accurate facsimile of what the data "looks like".

The next two sections present the specific approach taken and describe the experimental imagery produced. Appendix E describes details as well as other advanced considerations.

4.6.2.2 Specific Approach to Fixed Color Mapping

The following approach was used to produce color-stabilized imagery on the PFC. It represents a recommended procedure for evaluation as a method for producing supplemental or alternative image products.

Standardization of scenes against external effects, namely, solar geometry and atmospheric haze level, was accomplished by application of the spatially varying XSTAR algorithm [30, 36]. XSTAR corrects for differences in scene haze with good reliability. Without such corrections, the concept of a fixed mapping from data to color would make little sense.

The two primary components of the data, Brightness and Greenness, were used to drive independent dimensions of color variation. In this way the two data components can be independently assessed by the analyst.

The visual mechanism reads the dimensions separately, in a manner which is immediate and natural. Specifically, we used the data Brightness dimension to control color brightness on the film product, and we used the data Greenness dimension to control the degree of color saturation. The third dimension of color variation, namely hue, was effectively fixed since Tasseled Cap Yellow, which exhibits little variation, is projected in the direction of hue.

Control of film color was achieved through application of the modeling work of Juday. The model allowed us to specify color in the coordinates of a color space with certain desirable properties. In effect, we converted the PFC photographic system from one which was driven by projection of data onto color gun counts (which are nonlinear) to one which was driven by projection onto desired color space coordinates. The color space employed was L^*,a^*,b^* color space which is the standard uniform color space as designated by the CIE in 1977 (see Appendix E for some alternatives).

The color space coordinates give us control of color in two important respects. First, the coordinates separate color variation along its psychological dimensions. This allows us in principle to change hue, saturation and lightness independently of one another. The L^* coordinate controls color brightness while the a^*,b^* vector controls chromaticity -- its direction determines hue and its magnitude determines saturation. Second, the rate of change of color within the space is, in principle, uniform. That is to say Euclidean distance within the space, measured as $(L^{*2} + a^{*2} + b^{*2})^{1/2}$, is a consistent measure of the perceived difference between colors. When we map data into a uniform color space, we expect distance relationships in the data to be preserved in perceived color differences.

Colorimetrically controlled imagery was produced by mapping data variables linearly* into uniform color space coordinates. One can envision embedding the Tasseled-Cap Greenness/Brightness plane onto a plane in color space. We did this with the two constraints -- that data Brightness line up with L^*, thus controlling color brightness, and that data Greenness be perpendicular to Brightness, determining color saturation. The shape of the color space accessible to the PFC photographic system had a role in determining which hue we chose to fix upon. The shapes of two-dimensional slices of color space which have constant hue vary significantly throughout the space and we sought the one which best matched our data envelope. It turned out that the b* direction was a good choice for our needs. Figure 9 presents a stylized version of how we placed the data into the color space for our experiment. Our image produce aligns Greenness with c*, which fixes us on an orange or yellow-orange hue.

4.6.2.3 Observations and Conclusions

Experimental imagery and a color key were produced for the acquisition history of Landsat Segment 1663 in North Dakota throughout the 1977 growing season.

For all these images only one mapping from data to color was involved. We did not change the map from date to date by contrast stretching designed to enhance color contrast in the images. Thus, the analyst using such imagery can work with a single color key which can be explicitly shown to him.

---

*Note that one important advantage of mapping data to color is that the mapping need not be linear, yet the resultant projection into color is totally predictable.

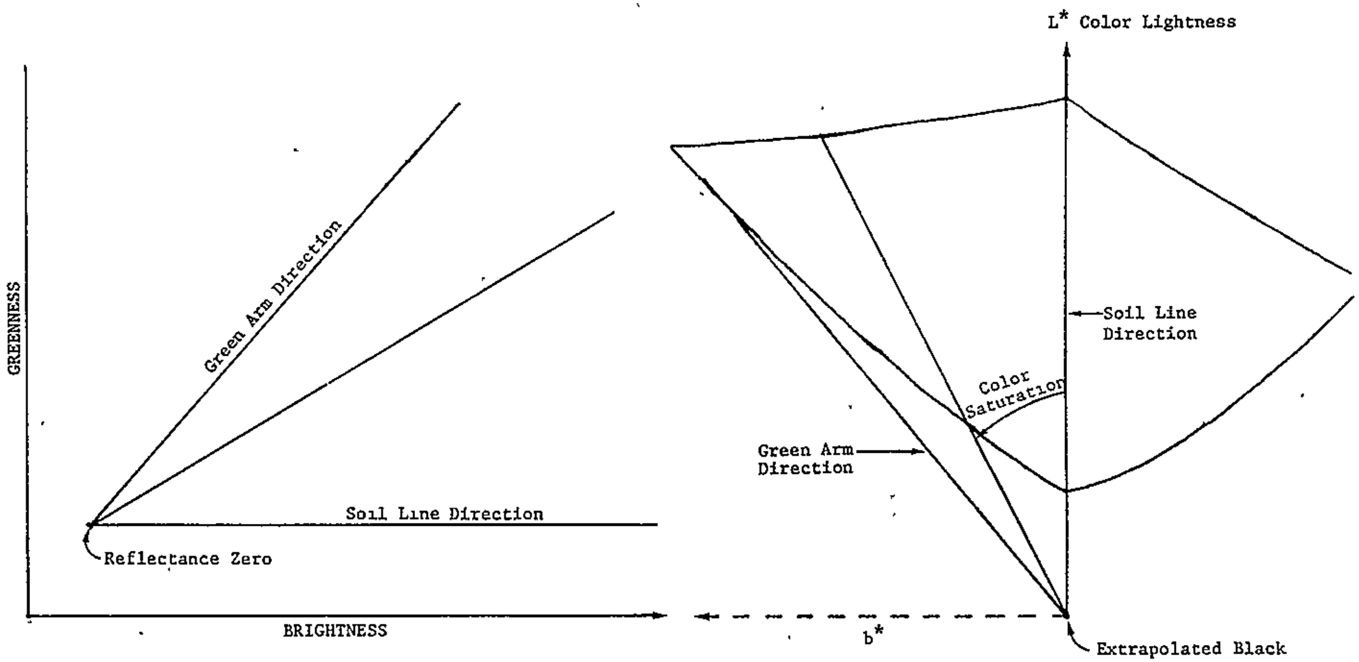


FIGURE 9. ADVANCED IMAGE PRODUCT

The imagery indicated contrast stretching for color resolution is in fact unnecessary. All color resolution was maintained regardless of the scene mean signal values. From an analytic point of view we have designed appropriate spectral resolution into our imagery. In Uniform Color Space we know the size of a color difference which is just noticeable to the human eye. We ensure that this mapping from data to color takes the smallest data unit, one Landsat count, into a color difference which is just noticeable. From an empirical point of view, these images compared favorably against currently employed products in the level of detail which can be discerned.

The extent to which advanced imagery can increase the ease and accuracy of image interpretation cannot be ascertained before it is tested by analyst interpreters. The major features of advanced imagery for Landsat--data standardization, use of principal components, mapping into Uniform Color Space coordinates, and a fixed color key--all aim to provide a consistent image environment within which interpretation can take place.

4.6.3 IMPROVED SPATIAL-SPECTRAL CLUSTERING ALGORITHM (SUPERB)

A desire to improve on performance capabilities of the BLOB spatial-spectral clustering algorithm arose from at least two needs. First was a need for a supervised mode to allow use of data in ground truth channels to form quasi-fields in situations where individual field boundaries were not encoded, i.e., where just the crop type was encoded. Field-to-field variability within a given crop type can be large. A second need arose out of results of the experiment in analyst-interpreter labeling of blob targets; as discussed in Section 4.9, the analysts were concerned about the presence of blobs having two or more disjoint parts. The fact of revision provided an opportunity to incorporate other desirable features and capabilities as well.

A development effort was conducted and the features listed in Table 8 were incorporated into a revised algorithm called "SUPERB", short for "SUPERvised Blob". Details of these developments are presented in Appendix D. SUPERB was used in preparing the data base described in Section 7 and Appendix D.

4.6.4 ESTIMATION OF CROP DEVELOPMENT STAGE

Uncertainties about the stage of crop development can contribute to analyst-interpreter labeling errors at two levels. The first level is between segments, the second within. Unknown to the analyst-interpreter, a given segment may be more or less advanced in development than others observed by Landsat on the same date. Even within a segment, individual fields of the same crop may be planted up to two or three weeks apart, causing confusion, especially in view of the 18-day repeat cycle of Landsat. Missing acquisitions just compound the problem. A need exists, therefore, for a capability to objectively estimate the crop calendars for individual fields (or quasi-fields) and individual pixels (or dots).

An investigation was initiated to work toward development of an objective method for estimating the crop calendars of individual fields, using multidate Landsat data, for use in labeling procedures. For this initial work, an empirical approach was taken. To date, limited examinations have been made of temporal-spectral Landsat data and field measurement data as a function of wheat development stage observed on the ground. Ultimately, we envision a crop calendar estimation methodology that will combine meteorological and spectral inputs, although this work was directed toward the spectral aspects.

TABLE 8. DEVELOPMENT HIGHLIGHTS OF SPATIAL-SPECTRAL CLUSTERING PERFORMED BY ALGORITHM SUPERB

Development	Significance
Ground truth supervision is available as an option.	Formation of quasi-fields can be forced consistent with ground truth boundaries. The resulting blobs can be used as ground truth fields, in some applications.
A new distance measure allows separate consideration of spatial and spectral effects.	It is easier to judge the significance of of threshold parameter values.
A pixel may be assigned only to a quasi-field associated with a neighboring pixel (or form a new quasi-field).	Quasi-fields now can be made to consist of a single group of adjacent pixels. Fewer distance calculations need to be performed.
When deciding whether to assign a pixel to a blob versus form a new quasi-field, pixels having the easiest decisions are processed first (within each scan line).	More candidate neighbor quasi-fields often become available to help improve later decisions. Biasing effects of sequential processing order are minimized.
Newly formed blobs are compared to their neighbor blobs, and combined if sufficiently similar.	If two blobs happen to start in the same homogeneous field, they may be combined.

The temporal profile of Tasseled-Cap Greenness and associated profile shift features were selected as the Landsat features for comparison with information on wheat development stage. The wheat development stages were measured on the Modified Feekes Scale. This scale is not linearly related to calendar date. Crop calendar shifts were calculated using spectral mean values for each field. Details of the analysis are presented in Appendix B, including example plots of development stage and Greenness vs. both acquisition date and shifted date.

This initial work in crop development stage estimation did not reach a conclusive stage. However, some issues were identified that could affect future work. It is recommended that ground observations of crop development stage in selected fields in segments be made more frequently than at 18-day intervals, even though spectral data may not be acquired more frequently. The sparseness of Landsat data, especially with missing acquisitions, will limit accuracy. The use of green profile shift technology appears helpful but improvements involving more complete characterization of profiles (e.g., Level 2 or Level 3 applications) and/or use of other spectral features, e.g., Brightness, may be required.

4.6.5 CONSIDERATIONS IN PROFILE FITTING

A procedure for fitting a spectral profile model to a set of multidate data points should include three steps: (a) data and model selection, (b) data normalization, and (c) parameter estimation.

The choice of data for fitting and the model form to be used depend on each other as well as on the level of use or application intended (c.f. Section 4.4) and the parameter estimation procedure. Important aspects of data are the type (e.g., pixel, blob or field mean, or segment mean), the quantity (e.g., selected pixels, all or selected fields, or a large fraction of pixels in segment), and the number of acquisitions

represented. The interactions of these and other factors in making a choice are addressed below and in Appendix A.

Data normalization is an important step that is often overlooked in data preparation. We recommend the use of normalizing transformations to reduce variability in spectral observations that results from factors other than crop type and condition. Normalization procedures are discussed in several places in this report (e.g., Section 5.2 and Appendix A).

Parameter estimation procedures are closely tied to the model form and the data aspects mentioned above. As discussed in Appendix A, use of time and/or Greenness offsets is important to obtaining good results with the exponential model forms we have developed (c.f. Section 4.4.3). Another important consideration discussed there is the choice of calculational procedure. Linear regression procedures have normally been used with these model forms after a logarithm transformation. A non-linear regression procedure is shown to give more accurate representation of peak Greenness values, at an increase in computational cost; the non-linear procedure is recommended wherever data volume does not make costs excessive.

Our experience and recommendations for parameter estimation are summarized next for selected applications at the first three levels of use (c.f. Section 4.4.2).

4.6.5.1 Level 1

Our recommended crop calendar shift calculation procedure for spring small grains is as described in Reference 3. It uses the model form of Equation 1 with a specified set of parameters as a standard reference profile. These parameters give a representation of spring small grains in the Northern U.S. Great Plains that appears

adqeuate for most shift calculations, given a minimum of three acquisitions, spaced no closer than 18 days, in the main part of the crop development cycle.

As a refinement of the procedure, we have on occasion utilized a two-step process. With it, a segment-specific profile is calculated after shifting all labeled small grains quasi-fields by amounts described by calculations using the standard profile. This new profile is then substituted for the standard one and the crop calendar shift calculation repeated. Differences in the two shifts calculated for each pixel have been slight, but use of the two-step process could be beneficial.

4.6.5.2 Level 2

A Level 2 application we have addressed is that of producing segment-specific estimates of the average peak Greenness value of all labeled spring small grains. These estimates serve as indicators of moisture stress for the Spring Wheat vs. Barley labeler discussed in Section 4.8 and Appendix G. For these estimates, we have used linear regression on spectral means for quasi-fields, after time shifts calculated using the standard profile.

4.6.5.3 Level 3

We explored profile fitting for Level 3 applications, where one desires good fits both before and after the day of peak Greenness for individual fields or quasi-fields. The original model form (Equation 1) showed deficiencies in fitting some situations, especially early in the season near the start of greenup. Therefore, after exploring other avenues, the model form of Equation 2 was developed. It still needs a Greenness offset (e.g., 25 counts) but does not use an initial time offset, since it falls exponentially from its value at the day of peak Greenness (which is determined in a separate Level 1 calculation).

Indications from a limited number of cases examined were that the overall performance of the second model form was better.

Another consideration is whether to use field means or individual pixels for fitting. No clear advantages were found for either. In both instances, edge pixels which may cross field boundaries should be eliminated from consideration. The procedure described herein works with at least three independent acquisitions during the growing season. However, more than three acquisition are desirable to permit a better representation of crop phenology.

4.7 SIGNATURE CHARACTERIZATION

The structure of the overall problem of objective labeling includes feature definition, signature characterization and labeling procedure development. In a working research environment these three categories flow into and merge with one another and it is not always clear where one begins and another ends. We will attempt a definition of signature characterization which allows these distinctions. We further wish to distinguish between signature characterization directed to the needs of analyst interpreters and signature characterization directed to the needs of machine labeling or area estimation. The following discussion is directed primarily to the problem of machine labeling.

4.7.1 OBJECTIVES OF SIGNATURE CHARACTERIZATION FOR MACHINE PROCESSING

It is the objective of signature characterization to provide a probabilistic relationship between the features observed, the collateral information provided, and the classes of objects which can give rise to those features under the given set of collateral conditions. This probabilistic relationship is expressed by the conditional likelihood function;

i.e., the probability density of the features, given the class observed and the conditions of observation.

$$L(y|c,v) \equiv Pr(y|c,v) \tag{3}$$

where

y = feature vector

c = class index

v = collateral condition vector

$L(y|c,v)$ contains all the information that could be available to a machine for classification or for area estimation. Given L, it would also be possible to determine the ultimate limits of accuracy to which area can be estimated, given a finite set of observations of the feature vector y. Thus, full knowledge of L would allow one to determine unambiguously whether a given set of features and collateral data was adequate to the required estimation task or whether new features or collateral data are required. If the feature set and collateral data set tested carries the full information content of the measuring instrument and procedures, then it could be determined whether a new measuring instrument is required.

The accurate estimation of the likelihood function is, however, not an easy task, for a variety of reasons. It requires both a large base of empirical data and accurate characterization of collateral features and conditions. We are not aware of an example in remote sensing applications where L has been accurately estimated and then exploited. It is our aim to come closer to such a characterization than has been achieved in the past.

In the context of area estimation for agricultural inventory using Landsat and collateral data, signature characterization has to satisfy applications criteria as shown in Table 9 and training criteria as shown in Table 10.

TABLE 9. APPLICATION REQUIREMENTS

Signatures applied must

- utilize available multitemporal acquisition histories
- accept a wide range of observation conditions
- represent accurately the likelihood functions of crop of interest and confusion crops for the available acquisition histories and observation conditions

TABLE 10. TRAINING REQUIREMENTS

The process of estimating signatures must

- utilize available multitemporal observations and interpolate to obtain time continuous signatures
- utilize observations representing a wide range of collateral conditions and interpolate to produce signatures continuous with respect to these conditions
- utilize a large number of randomly-drawn observation sets in order to insure inclusion of a robust variety of patterns

The first requirement in each of these tables emphasizes the problem of missing observations, due to the fact that Landsat passes over any given site only every 9 or 18 days and does not acquire data over every site on every pass primarily because of cloud cover. Further, in the presence of scattered clouds, observations may be missing for certain fields even when much of the segment is acquired. This is true separately for both signature training and for area estimation.

The second requirement in each table emphasizes the need for signature extension. It is known from LACIE experience that estimated signatures do vary significantly from segment to segment, partly due to physical processes dependent on collateral observables and partly due to random variation. The collateral variable dependence may be accounted for by training as a function of the collateral observations, if sufficiently well known. This results in a narrowing of the estimated likelihood functions for each condition of observation.

The third requirement emphasizes the need to include all of the random variation in the estimation of the likelihood function. This would enable use of the estimated likelihood function with confidence, both for making area estimates in an objective manner and for stating the error bounds on those estimates.

At this point an important question can be raised, "Given a procedure for training, with all attendant difficulties aside, how would one know that he had trained sufficiently in light of this last requirement?" The empirical, and qualitative, answer must be, "When marginal training no longer changes the estimation results, then one will know he has trained enough." In order to avoid some of the difficulties of a purely empirical approach, we can conceive of setting up a model which incorporates the training, signature estimation, and area estimation procedures. From a preliminary estimation of the signature, the model could then estimate the expected variance of area estimation both as a function

of the number of training samples and the number of survey or test samples, thus predicting the total number of training samples needed. Upon iteration with additional training samples, we would expect these predictions to converge to a satisfactory total number.

4.7.2 APPROACH

We have identified two approaches to satisfying the above requirements, distinguished primarily by the point at which feature extraction ends and signature characterization begins. The first of these is trajectory signature characterization, the second is generalized signature characterization. The definition and development of these approaches is discussed in Sections 4.7.3 and 4.7.4, respectively.

4.7.3 TRAJECTORY SIGNATURE CHARACTERIZATION

Trajectory signature characterization approaches the first two difficult requirements (missing data and collateral data dependence) by incorporating them into the process of feature definition, which is discussed earlier in Section 4.3. The general approach is to use physical modeling, field measurements, and Landsat data associated with ground truth data to establish definite mathematical forms defining the time trajectories of Landsat observables (e.g., Tasseled-Cap Greenness and Brightness) as a function of collateral observables for each class of interest and their confusion classes. In addition, there may be several free parameters, say ξ's, for each functional form. A good deal of preliminary training as well as physical insight goes into the establishment of these functional forms. Now, given a particular observation through time, $\{y(t_i)\}$, the values of the free parameters can be estimated and a measure of goodness of fit of the observations to the estimated curve $f(t)$ can be calculated.

At this point it would be possible to go directly to labeling; i.e., if the goodness of fit were best to a curve representing Class c, one would label the observation Class c. This approach is useful for obtaining insight into performance at a particular level of development of the feature definition approach; and if the feature definition supports near perfect labeling accuracy, the labels so derived could provide direct support to an area estimation scheme such as a stratified areal estimate. If however, labeling accuracy is not very high (i.e., is not >90%) it is likely to be necessary to employ a more sophisticated use of the trajectories, involving signature characterization.

An approach to characterizing the trajectory signatures is to treat the estimates of the free parameters, $\vec{\xi}$, and the goodness of fit measure, say r, as a derived set of features. Initially one could proceed using only trajectories for Greenness and Brightness for the crop of interest, say wheat. The task then is to estimate

$$\Pr(\xi | c, r)$$

for each class, conditional on the goodness of fit, r. The estimation of a conditional density function is discussed in the next section.

4.7.4 GENERALIZED SIGNATURE CHARACTERIZATION

Generalized signature characterization approaches the problems of missing data and collateral data dependence at the stage of defining the form of the likelihood function and training it, rather than at the feature extraction stage. In defining a statistical model form for the likelihood function, we have considered several possible approaches. These include a histogram approach, an expansion in terms of higher moments, and a multimodal Gaussian representation. In all cases, the way of dealing with collateral data dependence is the same and so we will discuss that first.

Figure 10 shows a hypothetical joint density function of spectral-temporal features, y, on the ordinate and collateral variables, v, on the abscissa, for a particular class, c. The density function, $Pr(y,v|c)$, is indicated by iso-density contours. In order to estimate this joint density in an unbiased way, it would be necessary to take an independent random sample of observations from Class c. It is, however, difficult to obtain an independent random sample of sufficient size from the domain of v. The distribution of v varies from year to year and is highly correlated from place to place within years. Hence an estimated probability density, $\hat{Pr}(y,v|c)$ has the peculiar property of being undersampled in the v domain while being sufficiently sampled in the y domain.

Consider the empirical density estimate $\hat{Pr}(v|c)$ for the same sample set. It is at least true that

$$\hat{Pr}(v|c) = \int \hat{Pr}(y,v|c)\, dy \tag{4}$$

By Bayes formula

$$\hat{Pr}(y|c,v) = \hat{Pr}(y,v|c)/\hat{Pr}(v|c) \tag{5}$$

This formula (Equation 5) holds for either the extreme case where all density functions are perfectly estimated, or the other extreme case where empirical density functions are represented by a discrete scatter plot or histogram. We make the assumption that this formula holds for the intermediate case in which density functions are represented by some functional form whose parameters are estimated from the data. Hence, Equation 5 provides the basis for a kind of regression to obtain a conditional density estimate of y given v. We proceed by estimating the joint density, computing the marginal density by integration and dividing to obtain the estimated conditional density.

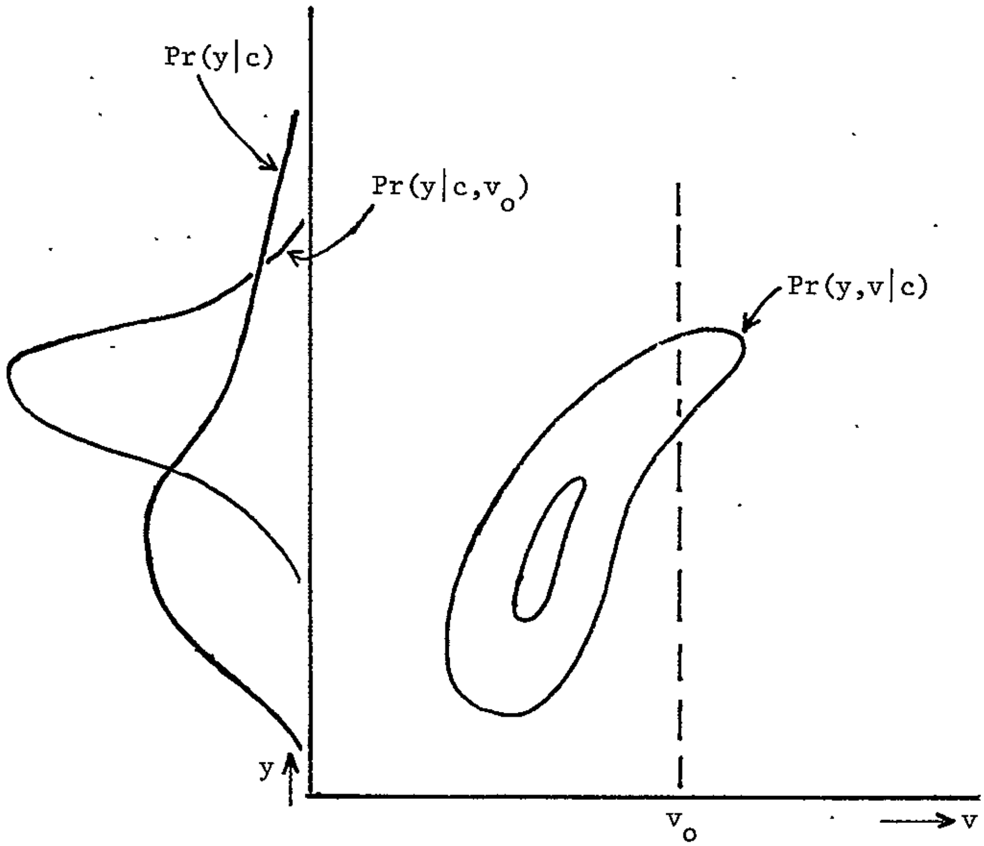


FIGURE 10. ILLUSTRATING THE CONCEPTS OF JOINT, MARGINAL, AND CONDITIONAL LIKELIHOOD FUNCTIONS

Further, if we use a Gaussian representation of the joint density then the indicated integration and division are analytical operations with matrix algebra, and this is true even if we represent the joint density as a convex mixture of several Gaussians, as shown in Appendix F.

A particular approach to the efficient representation of a density function as a convex mixture of Gaussians is the CLASSY algorithm [7]. We plan to adapt this algorithm for our use, to meet the requirements of missing data and of large feature vectors, as described in the following paragraphs.

The problem of missing data has been a central difficulty in all multitemporal signature extension efforts. The two aspects that concern us are: how to use incomplete data for estimation of crop area and how to use incomplete data for training. The solution to the first problem is conceptually straightforward. Suppose we have defined a complete vector y which consists of spectral features (say just Greenness) taken at interval closely spaced in time through the growing season, say every nine days. For winter wheat, this might be a 30-component vector,

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Suppose we actually have the likelihood function for this feature vector, $Pr(y|c)$. Let $\tilde{y}$ be an observation which we wish to classify, where $\tilde{y}$ contains observations at only a subset of times,

$$\{i\} \varepsilon \{1,\ldots,n\}$$

The the density of $\tilde{y}$ is the marginal of $Pr(y|c)$ integrated over the missing observations,

$$\Pr(\tilde{y}|c) = \int \ldots \int \Pr(y|c) \prod_{j \varepsilon \{i\}} dy_j \qquad (6)$$

Again, if $\Pr(y|c)$ is represented as a convex mixture of Gaussian densities, this integration is the sum of separate integrations for each component density and is a matrix algebra operation for each.

The problem of estimation of $\Pr(y|c)$ with missing data is more difficult. Boullion [31] has approached this problem by using an iterative procedure to make a maximum likelihood estimate of the density function given incomplete data. He does not assume any particular value for the missing components, but rather integrates over all likely values of the missing components using an estimate of their conditional density from previous iterations. However, this solution only considers single-mode Gaussian density functions, whereas we believe that multimodal density functions correspond more closely to reality. We have defined an approach in which the missing observations would be replaced by specific values, namely the conditional expectation based on previous iterations. This approach is explained in Appendix F. Briefly, this approach would involve running the CLASSY algorithm in a loop. At the end of each iteration CLASSY outputs a full-dimensional mean vector, a covariance matrix and a prior for each component density. From these, the conditional expectation of the missing data would be computed for each observation and the reconstructed data entered into the CLASSY algorithm for the next iteration.

There is another problem not mentioned above which has to do with the rather limited number of data points available for training, combined with the high dimensionality of the data vectors. This combination leads to near singularity in estimated covariance matrices which, in the CLASSY program, are inverted a large number of times. CLASSY, as it stands, makes no provision for inverting singular matrices (e.g.,

generalized inverse) and will have to be modified before it can be used with high-dimensional data. Since we expect a high degree of inter-temporal correlation in the data, it may be feasible to carry out all operations within CLASSY in a subspace of the full dimensional space. As a last step in each iteration of CLASSY, a full-dimensional estimate of the means and covariances would be calculated in order to compute conditional expectations of the missing data points for the next iteration. As yet, we have not tested this approach. Our plan calls for small scale experimental runs to determine the feasibility of the procedure described in Section F.1 of the appendix.

A word is in order as to why there are a limited number of data points. There are 23000 pixels in a LACIE segment and we have 50-100 ground truth segments to choose from. However, we are using quasi-field means instead of pixels, (since the major source of variation is field-to-field variation), and there may be only 400 of these per real class. The idea that 23000 pixels gives 23000 degrees of freedom for estimation is an illusion to begin with; the number of fields is a much better estimate of the number of degrees of freedom, albeit still too large. To substitute noisy pixel data in order to be able to invert matrices does not attack the heart of the problem, it merely masks it.

Trajectory signature characterization as described in Section 4.7.3 does not suffer from the difficulty of too few data points, since the number of features used is much smaller. Neither does it suffer from the missing data problem since it has already smoothed over missing data in the creation of features. Hence, CLASSY, or some other standard training approach, can be used directly on the derived features.

4.7.5 ISSUES RELATED TO SIGNATURE CHARACTERIZATION BY AND FOR ANALYST-INTERPRETERS

Analyst-interpreters develop their own characterizations of crop spectral signatures through years of training and experience. It is not easy to determine or define this characterization, especially since it also depends upon collateral information which may be available. Two issues are relevant to the investigation of objective techniques for labeling -- analyst labeling performance patterns and analyst extraction and quantification of collateral data.

In Section 4.9, we describe an experiment in analyst labeling of field-like targets and substantial analyses of the results. These analyses attempt to discover trends in the labeling performance that can be of value in subsequent developments. One analysis examines performance as a function of static spectral-temporal strata which in a sense could be descriptive of the analyst's conception of spectral signatures.

The second issue, analyst extraction and quantification of collateral data, was not addressed during this year. It is, however, an issue that will bear on optimization of the balance between man and machine in new labeling techniques.

4.8 REFINED MACHINE LABELER FOR SPRING WHEAT VS BARLEY

The problem of discriminating between spring wheat and other spring small grains, of which barley is the most important, emerged as one of the key problems in LACIE [1]. In response to that need, ERIM has conducted research to develop machine-based spring wheat/barley labeling techniques. Last year, a first-generation machine labeler was developed and tested. This work has been reported in detail [3] and is summarized in Appendix G of this report. Its distinctive features are its use of temporal-spectral profile fitting technology to estimate

crop calendar shifts for individual pixels and fields and its use of a distance measure in the Greenness vs. Brightness plane during grain ripening stages.

The test results indicated that the technique worked well in an area encompassing the segments used for developing the decision logic, but poorly in areas separated by some distance from those segments. When results were poor, most of the spring wheat was labeled as barley. It was also noted that segments with poor results were in geographic areas susceptible to drought conditions and likely to have brighter soils than the others. An effort therefore was conducted to refine the machine labeler, using the approach described in the next section.

4.8.1 APPROACH

Our approach to refining the labeler was keyed to developing and using an understanding of the biological and physical phenomena involved and their influences on the plant populations themselves. A sequence of five steps was employed, as detailed in Appendix G. The first was to formulate hypotheses to help establish a set of potentially useful directions to pursue. Second, pertinent physiological relationships and effects on spectral response were identified. Third, crop canopy reflectance modeling was conducted to predict responses. Fourth, field measurement data were analyzed. Finally, Landsat data were analyzed to learn if they supported or refuted the hypotheses and to provide the specific algorithm parameters.

4.8.2 RESULTS OF ANALYSIS

Hypotheses were established to the effect that moisture stress and soil brightness were likely causes of altered spectral signatures. Literature survey showed that effects of prolonged moisture stress include reduction in plant height, less tillering, thinner leaves, and an increased rate of plant development.

Reflectance modeling was performed and used to predict effects of both moisture stress and soil brightness. Both were shown to exert detectable influences on wheat canopy reflectances and, perhaps most importantly, on the distance measure used for discrimination, during the time period of greatest separability. The greatest effect of moisture stress was to cause an apparent advance in the time period of maximum spring wheat/barley separability. The effects of soil brightness were evident only when canopy closure was not full, i.e., as in early or late season or when canopies are stressed. These soil brightness effects were an offset or increase in the distance values and a changed slope during the time period of maximum separability.

Another key result was the observation that the moisture stress and soil brightness factors should be distinguishable and detectable by separate processes. Moisture stress should be detectable in a reduced Greenness peak for small grains and soil brightness should be measurable in early season acquisitions.

Landsat provided the major source of data for the specific labeler refinement. Two iterations of data analysis were carried out. The first iteration operated on pixels of small grains from six segments. It employed a new two-step refinement of the original crop calendar shift estimation technique. The added step employs a segment-specific reference profile, based on the first-step shifts, to compute refined crop calendar shift estimates. Also, the peak Greenness value of this segment-specific reference profile was used as an indicator of segment moisture stress. The apparent time period for separability of wheat from barley was determined through a combination of quantitative (e.g, discriminant analysis) and qualitative methods. These results supported the hypothesized relationships and a preliminary revision of the labeling logic was devised and tested, with encouraging results.

The second iteration utilized pixel data from the same six segments plus a seventh, together with the new profile model (Equation 2 of Section 4.4 and Appendix A) in the second step of the segment-specific shifting procedure. See Appendix G for details of the analysis performed. Data processing and availability problems precluded use of substantially more segments, as had been originally planned. Relationships consistent with the hypotheses and our physical understanding of the situation were developed although not as precisely as desired. Nevertheless, a refined decision logic was specified in a procedural form for test and evaluation.

4.8.3 REFINED DECISION LOGIC

The refined labeling logic allows for potential adjustments to three parameters that describe the original decision line, based on segment-specific spectral indicators. The available Landsat data were sufficient to support specification of only two of the three; these are the starting day of the period of separability and the initial distance value. The third parameter, the slope of the decision line, was fixed at the slope of the original decision line.

The key elements of the refined decision rule are illustrated in Figure 11. Note that normal, non-stressed canopies have peak Greenness values greater than G*, and that no adjustment in starting day is needed for them. Also, they have a good canopy closure so soil brightness adjustments are not required for them either. As moisture stress increases, peak Greenness decreases and the starting day moves to an earlier day. As soil Brightness increases under stressed canopies, the decision line move up. Appendix G contains specific descriptions of these decision rules, discusses their physical interpretation to a greater extent, and specifies a procedure for using them in labeling.

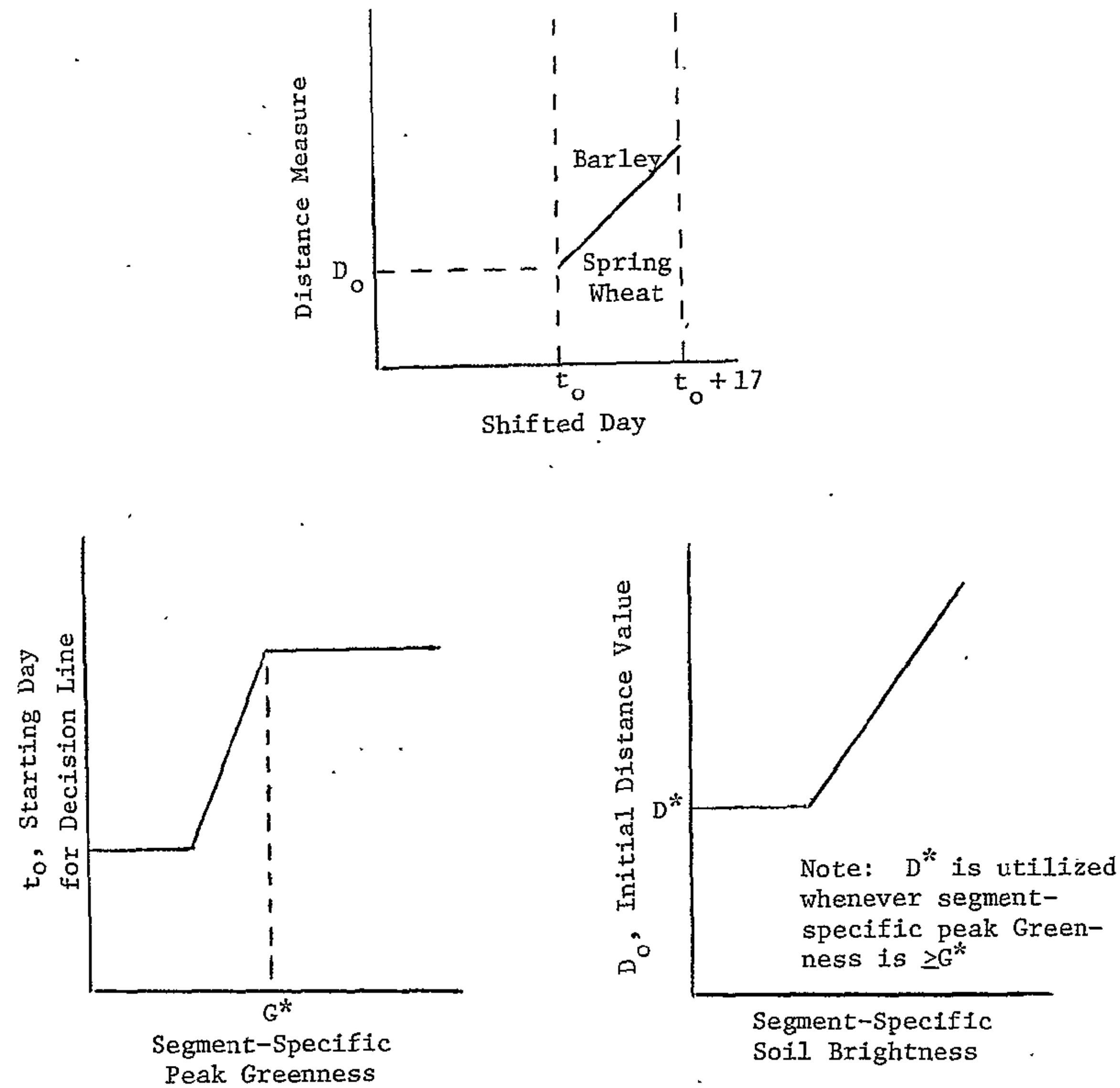


FIGURE 11. ILLUSTRATION OF REFINED DECISION RULE

4.9 A STUDY OF ANALYST-INTERPRETER LABELING OF FIELD-LIKE TARGETS

Ultimately the performance of a crop inventory system is dependent upon the accuracy of its measurement sources. Critical to most stratified area estimation procedures is the performance of analyst-interpreters who provide labels for sampled targets.

Image interpretation was extensively used in the Large Area Crop Inventory Experiment. In the earliest phase of LACIE, targets were analyst-defined fields. Later, Procedure 1 utilized systematically distributed 'dots' as targets. This section documents the first study made of using machine-defined quasi-fields or blobs as labeling targets. Appendix C describes the experiment design and provides details related to the experiment. This section describes the objectives of the experiment, the procedure utilized to label targets, and results of evaluations of the analyst performance.

4.9.1 OBJECTIVES OF THE ANALYST-INTERPRETER BLOB LABELING EXPERIMENT

The overall objective of this experiment was to gain an understanding of the analyst-interpreter labeling process in a field-labeling environment. It was hoped that insight would be gained in how individual analysts apply internalized decision-making processes to discriminate small grains from other canopies.

Major considerations in this study included:

(1) Developing and examining procedures for the use of field-like targets for labeling.

(2) Evaluating individual analyst performance in labeling these targets.

(3) Evaluating between-analyst consistency and its implications for improving labeling performance.

(4) Evaluating the impact of analyst performance on consistency and accuracy of crop proportion estimates.

(5) Establishing a data base for evaluating Procedure M (Section 6).

Specific objectives related to these major considerations are listed in Table 11.

It should be emphasized that major consideration is given to understanding the performance of analysts working individually without consultation. This approach was taken, rather than the Phase III LACIE labeling situation wherein analyst-interpreters consulted. This was done in order to gain an understanding of the decision-making process of individuals and how that process can vary from analyst to analyst.

4.9.2 THE LABELING PROCEDURE

Blob targets from 18 LACIE TY78 (Transition Year 1978) blind sites from the Northern Great Plains were labeled <u>grain</u> or <u>non-grain</u> independently by each of three LACIE-experienced analysts. To insure uniformity of the labeling procedure, that procedure was defined jointly by the analysts and their supervisor prior to the actual labeling. In addition, each analyst labeled the segments in a different random order, repeating the first three at the end of the sequence in order for us to examine learning trends (see Appendix C for details).

The procedural steps employed are summarized in Table 12. These procedures are not operationally recommended, but were defined for use in this experiment alone. For example, all blobs with interior pixels were labeled whereas operationally only a subset (e.g., 40 or 100) would be labeled. Any deviations from the standard procedure were noted by each analyst on a segment comment form. No substantial deviations were recorded. These notes are summarized in Appendix C.

TABLE 11. SPECIFIC OBJECTIVES OF ANALYST BLOB LABELING EXPERIMENT

- Develop and examine procedures for the use of field-like targets for labeling
  - Develop labeling aids and procedures
  - Analysts subjectively evaluate BLOB labeling approach

- Evaluate individual analyst performance in labeling targets as a function of:
  - Target level variables
    - related to target structure
      - -- blob size
      - -- blob purity
    - related to target state.
      - -- crop type
      - -- acquisition history
      - -- crop calendar (profile diagnostics)
      - -- crop condition (peak green)
  - Spectral strata (examine as a function of static spectral trajectory zones)
  - Segment-level variables
    - percent agriculture
    - percent grain
    - percent confusion crop

- Evaluate joint or between-analyst performance in labeling targets
  - Analyst consistency
    - analyst learning trends
    - absolute between-analyst consistency
    - consistency vs. accuracy
  - Average, vote and consensus labeling performance compared to individual performance
  - Evaluation of low confidence labels

- Evaluate analyst performance and consistency in estimating crop proportions

- Establish a data base of AI labels for evaluating Procedure M.

TABLE 12. LABELING PROCEDURE FOR PROCEDURE M TEST*

The procedural steps for labeling BLOBS (clusters) for the ERIM Procedure M test are outlined in this document.

1. Each analyst will label the BLOBS in the 18 TY segments (in a pre-specified random order).

2. Labeling will be performed independently by each analyst.

3. Each analyst will relabel the first three segments after all 18 segments have been labeled.

4. After labeling each segment the analyst will fill out the "segment comment forms".

5. After labeling all 18 segments and prior to relabeling the first three segments the analyst will complete the "final comment form".

6. Detailed BLOB labeling instructions are given below:

   a. For each segment, BLOB overlays keyed to the LACIE sample segment products (Products 1, 2 and 3) have been generated as well as line printer maps of the BLOBS for each segment. The line printer maps will be used by the analyst to record the label for each BLOB.

   b. A BLOB is to be labeled spring small grains if it is at least 50% spring small grains, labeled non-spring small grains if it is less than 50% spring small grains. When there is considerable question about which label to assign a BLOB, the BLOB will be labeled based upon the analyst's best guess and so designated. If a BLOB is mixed (i.e., composed of approximately 50% spring small grains and non-spring small grains) it will be flagged and labeled on the line printer map.

   c. The label for each category will be coded on the line printer map as follows:

      - Spring Small Grains - Red
      - Non-Spring Small Grains - Green
      - Questionable Spring Small Grains - Question Mark (?) on Red BLOB

*These procedures were developed by three LACIE analysts under the supervision of Mr. Robert Payne of Lockheed Electronics Corporation.

TABLE 12. LABELING PROCEDURE FOR PROCEDURE M TEST* (Cont'd)

- Questionable Non-Spring Small Grains - Question Mark (?) on Green BLOB
- Mixed BLOB - The BLOB will be labeled using the >50% criterion and then outlined in blue.

d. The analyst should examine the acquisition listing in the table in order to become familiar with the acquisitions used for generating the BLOBS.

e. Grid overlays (10 pixels by 10 scan lines) will be keyed to the line printer map in order to facilitate analysis.

f. Each of the 18 segments will have a packet consisting of the same material (ancillary data, maps, normal crop calendars, crop calendar adjustments, etc.) used during LACIE TY operations. The exception to this is that the spectral aids will not be used for crop identification (because they do not match the blob targets).

g. All analysis techniques (i.e., image interpretation techniques will be the same as those used during the LACIE TY operations).

NOTE: It is anticipated that the time required to label and complete the evaluation forms will be approximately 10 hours** (per segment).

*These procedures were developed by three LACIE analysts under the supervision of Mr. Robert Payne of Lockheed Electronics Corporation.

**Actual time averaged approximately eight hours.

It should be noted that specific decision criteria for labeling blobs were not included in this procedure. It was expected that the decision criteria developed in labeling dots as grain/non-grain through years of LACIE experience would form the base for the analysts' decision-making processes. It was, in fact, variation in each analyst's perception of what was a 'grain signature' that was expected to lead to variation in labels. Due to the short time frame available for organizing and conducting this experiment, spectral aids in the form of trajectories in Greenness vs. Brightness space were not developed for the blobs; however, all other standard LACIE ancillary data were available and recommended for use as usual.

4.9.3 ANALYST RESPONSE

Three analysts labeled an average of 379 blob targets for each of 18 segments over a six-week period of time, averaging eight hours effort for each segment. Their diligent effort is documented in the 'Report on the Analyst Test Using the ERIM Blob Labeling Procedures' [32]. The analysts' reactions to the overall blob labeling process provide an insightful critique of the blob labeling environment. Table 13 is extracted from their report and summarizes significant findings and recommendations. The analysts' reaction to the blob targets, in particular with regard to disjointed blobs, was a motivating influence in the development of SUPERB, a field-finding algorithm that insures spatial contiguity (refer to Section 4.6.3 and Appendix D). In addition, their reaction to small blobs helped motivate one estimation approach (described in Section 5.5) which utilizes 'extended' labels. It is expected that the definition of operational procedures will call heavily upon analyst recommendations, in particular the selection of appropriate acquisitions for blobbing.

TABLE 13. THE SIGNIFICANT FINDINGS AND RECOMMENDATIONS CONTAINED IN THE ANALYST BLOB LABELING REPORT

1. Strong Points

- BLOBS are easier to label than the dots used in Procedure 1 (P-1). A Blob represents a field center and does not contain border or edge dots as may a P-1 dot.

- Blobs represent field centers rather well.

- If ERIM reduces the number of Blobs to approximately 100, as currently planned, labeling of Blobs should be as efficient or perhaps more efficient than the labeling of dots in Procedure 1.

2. Problem Areas

- The Blobbing technique, as currently implemented, produces too many blobs for labeling (400-600).

- Small or stripped fields do not blob or cluster very well.

- Acquisitions selected for Blobbing by ERIM for the labeling test were not always optimum.

- Blobs were frequently disjointed which resulted in labeling difficulties.

- Small Blobs containing only 1 to 3 pixels are difficult to label.

3. Recommendations

- Acquisition selection for Blobbing should be based on multitemporal spectral information as well as spatial information (spatial data was used as the primary selection criteria for the test).

- Research should be conducted into the small fields problem.

- Reduce the use of single-pixel blobs whenever possible.

- Modify the line printer blob map and Production Film Converter (PFC) blob overlay. The current format of these two products is not conductive to efficient analyst labeling.

In addition to their report, analyst comment forms filled out upon completion of labeling in each segment and an overall summary after labeling all segments provided information with regard to critical aspects of their labeling effort. These forms are summarized in Appendix C for the questions listed in Tables 14 and 15.

4.9.4 EVALUATION OF ANALYST-INTERPRETER PERFORMANCE

Labels from each of the three analysts (referred to as Green, Red, and Blue) for over 6000 blob targets in 17 LACIE TY78* blind sites were incorporated into an SPSS-formatted (Statistical Package for the Social Sciences) data base. Evaluation of analyst performance was conducted for each of the specific objectives listed in Table 11. Appendix C describes the experiment design, the data base, and analyses in detail. This section will summarize major results of the experiment.

Evaluation of analyst-provided labels was carried out by examining each analyst's performance individually. In addition, two types of labels were fabricated from the three labels. These were a 'vote' label and an 'average' label. A vote label is assigned by the majority label of each blob (i.e., if at least two of the three analysts labeled it grain, the blob is assigned a grain label). The average label simply averages the three labels (e.g., if two analysts label grain, the blob is two-thirds grain).

Table 16 presents the overall 17-segment accuracies in terms of the percent of blobs correctly labeled. These results are similar to those experienced in Procedure 1 dot labeling. The wide variation from analyst to analyst, especially in labeling grains is noteworthy. Vote and average labels were used as two techniques to evaluate the joint use of analyst labels. Since patterns detected in individual analyst labels were also detected in the vote label, it will be used to illustrate analyses.

*Ground truth unavailable for one of the original 18 segments.

TABLE 14. QUESTIONS DIRECTED TO ANALYSTS UPON COMPLETION OF LABELING FOR EACH SEGMENT

- Were Landsat acquisitions deficient?
- Was choice of acquisitions for BLOB optimal?
- Do blob interiors seem pure?
- Do blob patterns match field patterns?
- Did you have to change your procedure?
- Describe:
    - percent agriculture
    - percent small grains
    - major grain crops
    - average field size
    - topography
    - apparent moisture

TABLE 15. QUESTIONS DIRECTED TO ANALYSTS UPON COMPLETION OF ALL LABELING

- Describe general impressions of labeling blobs
- List major problems and strong points
- Comment on suitability of analyst aids
- What procedures did you develop to organize and use the product?
- Recommend improvements

TABLE 16. PERCENT OF BLOBS* CORRECTLY LABELED
IN 17 LACIE TY78 BLIND SITES

	Percent Correct	
Label	Grain	Not Grain
Green	68.6	93.9
Red	71.1	89.8
Blue	51.0	95.3
Vote	66.8	94.9
Average	63.6	93.0

*For this chart a blob is considered grain if at least 50% of it is grain.

Table 17 presents overall proportion estimates derived directly from the analyst labels and compared to the true estimate* for the labeled targets. Each estimates the proportion of grain. The vote, average, and Red labelers produced the best estimates in an RMS sense, and the Red labeler provided least overall bias.

Analyses led to a number of observations summarized in Table 18. These will be discussed in the next section with reference to the experiment objectives.

4.9.4.1 Analyst Performance Related to Target, Spectral and Segment Level Variables

Target-Level Variables: A number of parameters related to the structural and agronomic makeup of blobs were related to analyst performance. These included blob size, blob purity, crop type, crop calendar, and crop condition. A significant relationship is expected between the availability of appropriate spectral time history and labeling accuracy; however, this experiment was not designed to address this issue as segments were selected to minimize the impact of inadequate acquisition history.

The target structure -- its size and purity** -- was found to be related to analyst performance. For the most part, blob targets were pure with respect to crop type. Though analysts were instructed to use 50% purity levels as a decision threshold, Figure 12 reveals a decision behavior that is logistic in nature. A blob that was half grain would be labeled as grain about 35% of the time, which corresponds to half the grain labeling accuracy for pure grain blobs.

*This excludes blobs not labeled by analysts.

**Since blobs are machine-defined, they may not be strictly pure but may overlap grain and non-grain fields. Blob purity is discussed in Section 6.2.

TABLE 17. GRAIN PROPORTION ESTIMATES DERIVED FROM ANALYST LABELS FOR ALL BIG BLOBS IN 17 LACIE TY78 BLIND SITES

Label	Proportion*	Bias	Std. Dev.	RMS
Green	.297	-.038	.110	.116
Red	.326	-.010	.099	.100
Blue	.231	-.104	.081	.137
Vote	.285	-.050	.089	.102
Average	.283	-.051	.084	.098
Truth**	.335			

*Blobs were weighted by their sizes.

**Includes only those blobs labeled by analysts.

TABLE 18. FINDINGS OF THE ANALYST LABELING EXPERIMENT

- Non-grain labeling accuracy is significantly greater than grain labeling accuracy, indicating a tendency toward conservative labeling of grains.
- Spring wheat was labeled correctly as grain more consistently and more often than barley and oats (oats were labeled as grain at a 30% rate in North Dakota, but at a 90% rate in Minnesota).
- Summer crops and grasses were not a significant problem with respect to errors of commission.
- Significant between-segment variations in the labeling of spring wheat, barley and oats were not found to be related to the proportions of grains or agriculture present in the segment.
- The probability of a given label appears functionally related to the target purity.
- Analyst accuracy increases as target size increases.
- Grains are mislabeled if they do not conform to a 'standard' spectral profile.
  - Correctly labeled grains exhibit similar spectral profiles across all segments
  - Incorrectly labeled grain trajectories differ distinctly
  - Linear discriminant analysis of crops at a segment level revealed accuracy patterns similar to analysts
  - A standard wheat signature concept may be key to analyst labeling behavior
- Distinctive relationships were observed between labeling accuracy of grains and profile derivates, shift and peak Greenness.
  - Error increases as shift deviates early or late from the segment norm
  - Error increases when peak green is low
  - Errors arising from late shift are not as prominant in the presence of high peak green values

TABLE 18. FINDINGS OF THE ANALYST LABELING EXPERIMENT (Cont.)

- The use of a vote or average label based on three independent analyst labels provides a more stable estimate than one based on the selection at random of any one analyst, and more accurate than the average label of two analysts.

- Proportion estimates based on each analyst's labels underestimated wheat and were statistically described in a linear regression as primarily offsets.

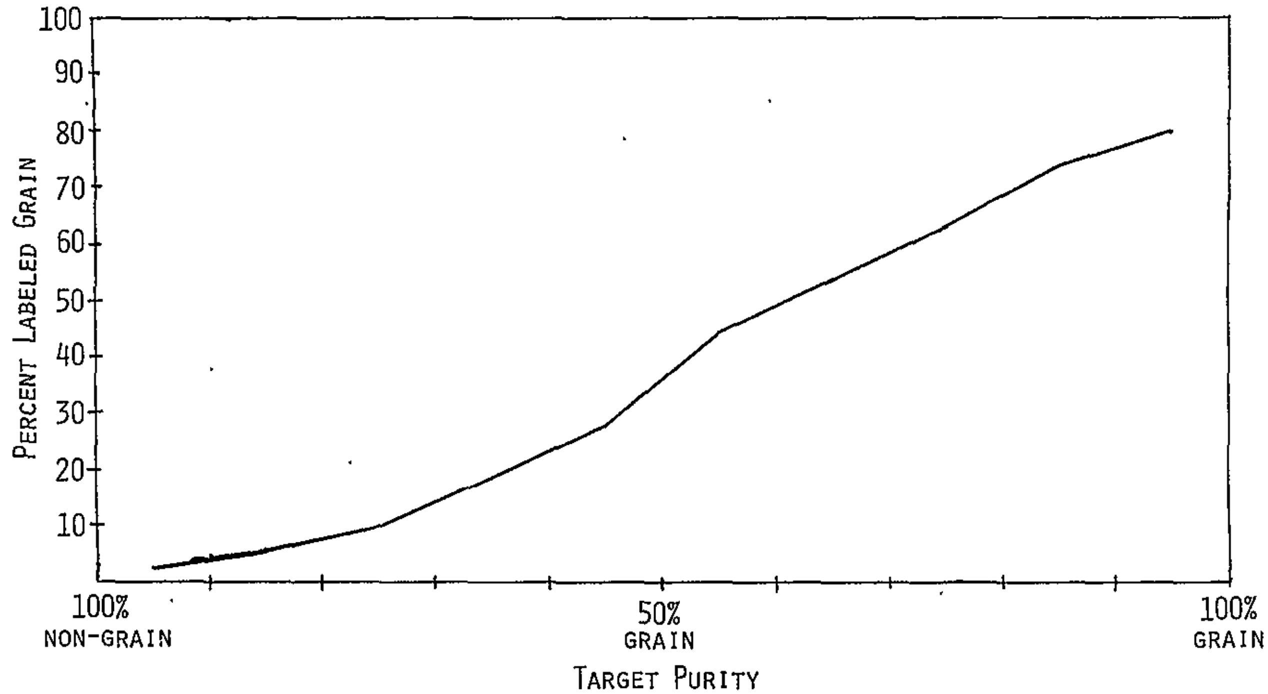


FIGURE 12. ANALYST VOTE LABELING ACCURACY AS A FUNCTION OF TARGET PURITY*

*OATS WERE EXCLUDED FROM THIS GRAPH SINCE THE LOW LABELING ACCURACY OF OATS AS GRAIN IS LIKELY DUE TO OTHER CAUSES.

Labeling accuracy was found to be correlated to the size of the labeling target. Figures 13 and 14 illustrate the relationship detected for each analyst and the vote label. Note that accuracy improves whether the blob is grain or non-grain. This tendency was found significant in each analyst and in the vote label for both the Pearson test and the Spearman rank test. It is conjectured that accuracy in labeling dots in Procedure 1 is similarly related to the size of the field in which the dot appears.

As is elaborated upon in Appendix C, range land and non-crop land were not confused as wheat. It is particularly interesting to note that in North Dakota, oats where mislabeled at a 70% rate, yet oat fields were generally large (contradicting the field size trend). These oat fields were generally earlier, hence may be confused with pasture. Yet, accuracy in Minnesota was over 90%.

Since analysts were not instructed to label non-grains by crop, errors of grain omission, which were prevalent, cannot be related to the presence of specific confusion crops. No significant relationship was found relating analyst accuracy (in terms of percent correct classification, or proportion estimate) to the proportion of wheat or grains. However, Figure 15 illustrates an unexpected trend in relating error to other canopies. All three analysts displayed a slight tendency, represented here by the vote label, to greater accuracy in labeling grain as grain when crop land other than grains was present. Similarly, wheat was labeled as grain more often in the presence of non-wheat canopies. Since the prevalent non-crop canopy is range land (including pasture and other grasses), it is hypothesized that confustion of grains as range land may occur.

In order to assess the relationship between labeling accuracy and agronomically related conditions, the profile diagnostics, crop calendar shift and peak Greenness were computed for each blob, as described in Section 4.4 and Appendix A. It was found that blobs whose estimated crop calendar shift deviated greatly, either early

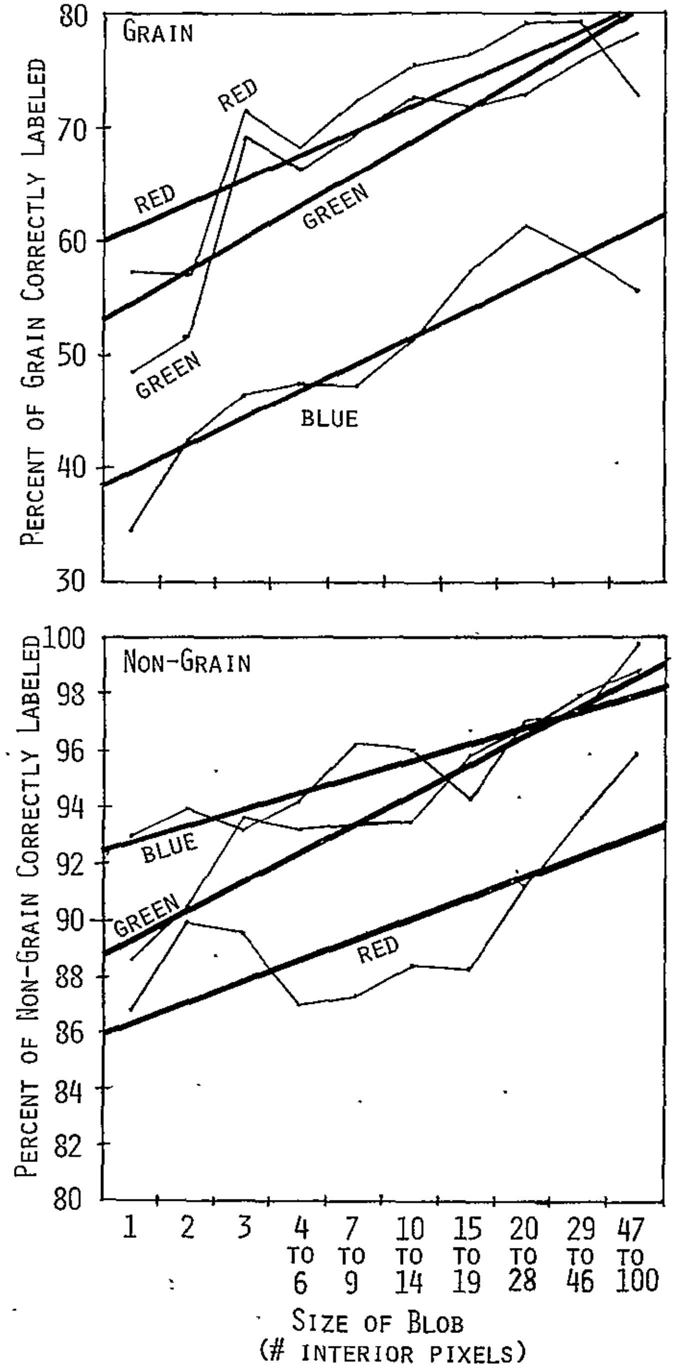
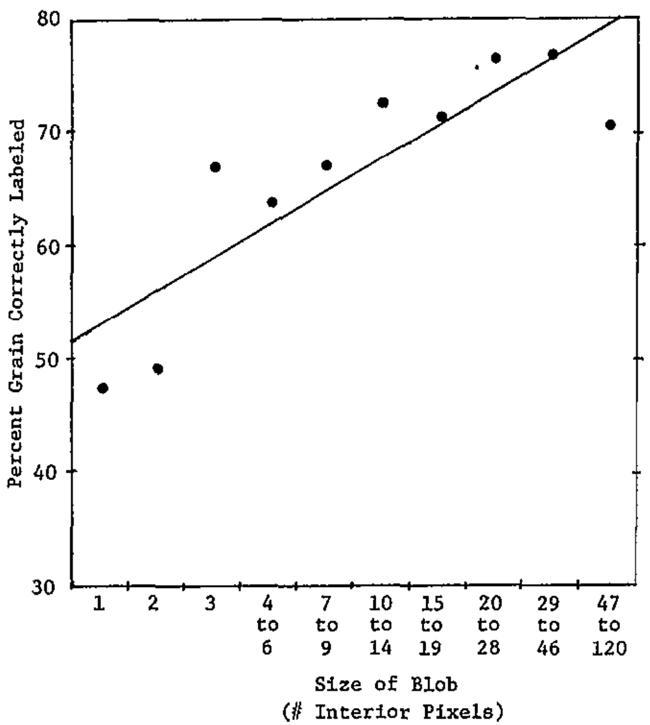
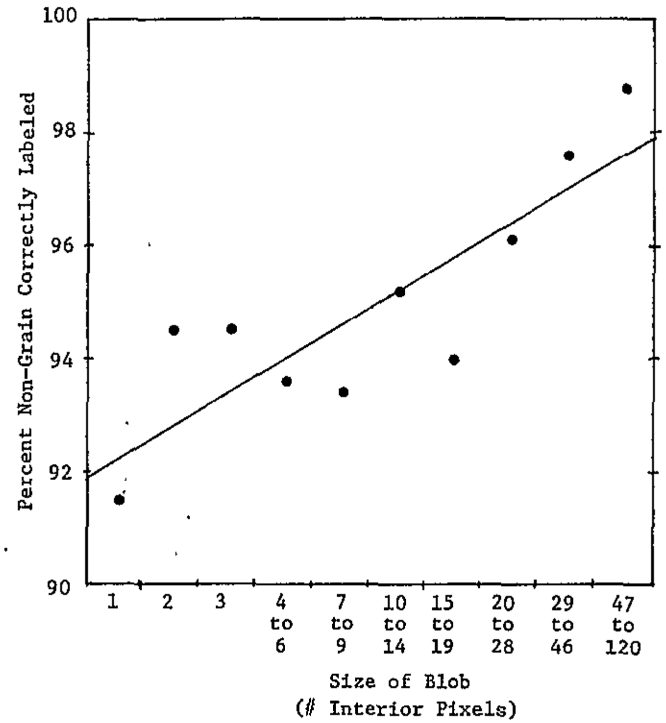


FIGURE 13. DEPENDENCE OF ANALYST ACCURACY ON TARGET SIZE

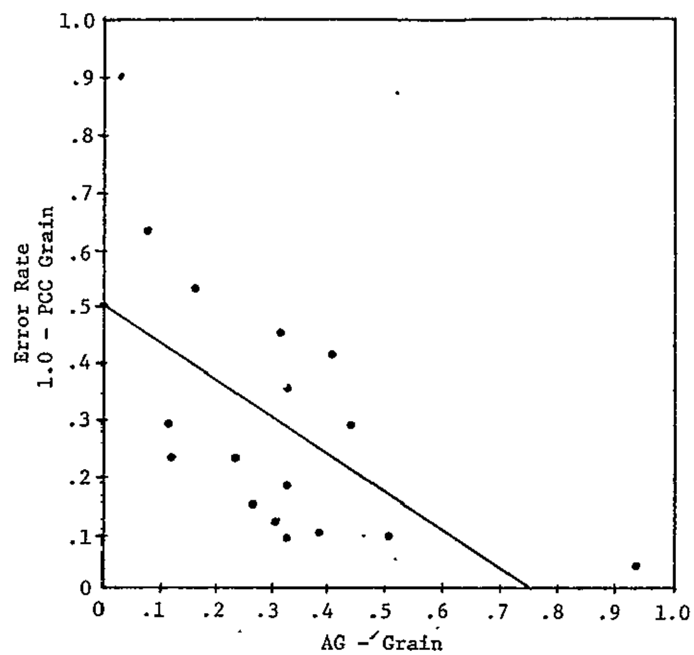


(a) Grain Labeling Accuracy by Vote vs. Target Size



(b) Non-Grain Labeling Accuracy by Vote vs. Target Size

FIGURE 14. DEPENDENCE OF ANALYST VOTE ACCURACY ON TARGET SIZE



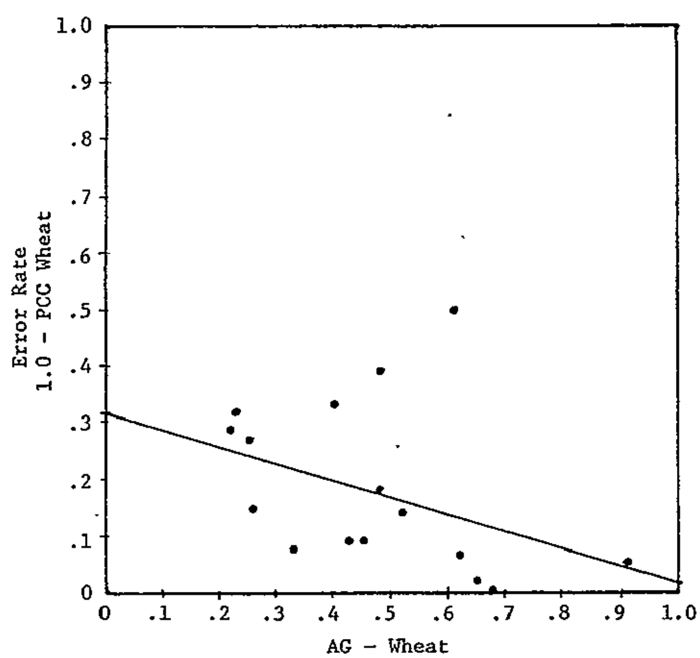


FIGURE 15. DEPENDENCE OF PERFORMANCE ON THE PRESENCE OF NON-GRAIN CANOPIES

or late, from the segment-specific norm were less accurately labeled. Indeed, even segments whose average shift deviated from a norm among segments achieved lower accuracy. When estimated peak Greenness was very low, both early and late shifts correlated with poor performance. However, errors for high peak Greenness levels were found only with early shifts. These findings indicate that the agronomic state of a target is an important influencing factor in labeling performance.

Spectral Strata: The dependency of analyst performance as a function of the various crop signatures manifested is a key issue in understanding labeling behavior. It was previously suggested that substantial variations of that signature from a 'normal' state due to variations in crop condition or in crop calendar resulted in reduced accuracy. To evaluate performance as a function of the spectral crop signature, an analyses was initiated that employed a physically-based temporal/spectral stratification technique that divides spectral patterns in time in a manner that permits a standard physical interpretation. This technique is described in Section 6.2.

Analysis to this point has revealed analyst preferences toward certain signature patterns over others that are equally likely. A simple illustration is provided in this section. Figure 16 illustrates a potential spectral path grain could follow through time. The path illustrated conforms to 'normal' crop development pattern wherein a potential grain target would exhibit a bare soil signature initially that geeens up to a peak value and senesces, during which the grain signature would brighten and decrease in Greeness. Considering a second pattern which is the complement of this, we are able to examine analyst performance as a function of each spectral path. No correction is made to account for segment-specific crop calendars.

As is seen in Table 19, although nearly 40% of the grain spectral trajectories do not follow the stated 'normal' path, they are labeled at an average rate of 57.2% correct, where the 60% in the first path are accurately labeled at a rate of 82.2% correct.

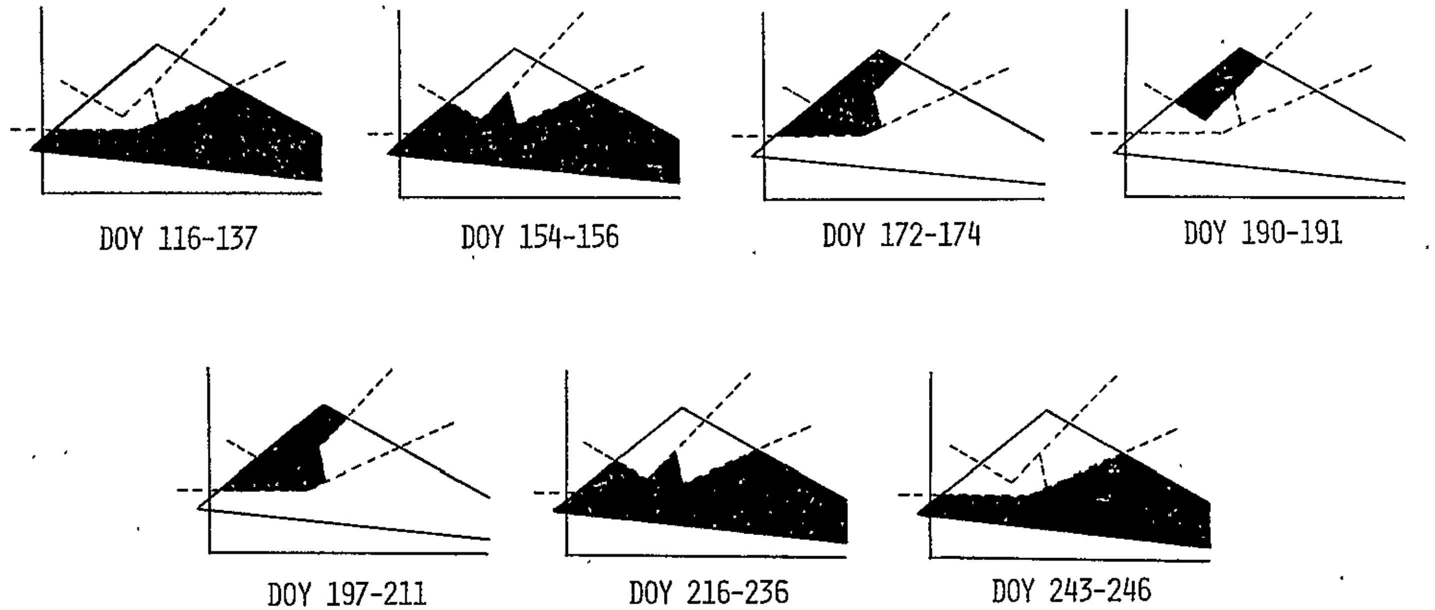


FIGURE 16. TEMPORAL-SPECTRAL TRAJECTORY PATH FOR STRATIFICATION OF ANALYST PERFORMANCE (FOLLOWED BY 60% OF GRAIN)

TABLE 19. ANALYST ACCURACY RELATED TO TEMPORAL-SPECTRAL PATH OF BLOB MEANS.

	Percent of Blobs That Follow Path			Percent Correctly Labeled									
				By Vote		Red		Green		Blue		Average	
Path	NG	G	G/(NG+G)	NG	G	NG	G	NG	G	NG	G	NG	G
'Normal Grain'	12.4	60.4	83.0	91.2	85.5	90.2	85.6	85.6	88.6	93.3	72.3	89.7	82.2
All Others	87.6	39.6	31.1	98.2	59.8	97.6	61.3	93.3	66.7	97.7	43.6	96.2	57.2

4.9.4.2 Between Analyst Labeling Performance

Analysis of analyst agreement on labels provides insight to factors influencing analyst performance. Table 20(a) displays the percent of analyst decisions in correct agreement in labeling various canopies as grain or non-grain. Note that 19.9% of all grain blobs were incorrectly labeled by all three analysts. If this were strictly a random event dependent on individual analyst labeling performance, one would expect only 2.8% of the targets to be missed by all analysts (Table 20(b)). This is a very clear indication that the pattern of labeling behavior is being consistently influenced by external factors.

Table 21 illustrates the probability of a correct label as a function of analyst agreement. This provides an indication of the confidence one can have in the analysts' labels when regarded jointly. For example, this table illustrates that a non-grain label from all three analysts was truly a non-grain canopy 98.0% of the time. Whereas, when a dissenter appeared among the three analysts, the two in agreement were correct only 69.9% of the time. The correctness of a grain label cannot be viewed as confidently, even if all three analysts agree upon the label. Information is carried in the dissenting label.

4.9.4.3 Analyst Performance in Estimating Crop Proportions

Two issues relate to analyst performance in estimating crop proportions. First, do estimates based on aggregating all analyst labels relate to the true proportions? Secondly, are analysts in agreement with one another even if proportion accuracy is poor?

Earlier, Table 17 provided proportion estimation performance for each type of label. Figure 17 illustrates the relationship between estimates based on the vote label and ground truth. The tendency is to underestimate grains, but in a consistent manner. The R = 0.83 and a regression line determined primarily by offset both indicate a consistent trend toward underestimating the presence of grains.

TABLE 20. ANALYST CONSISTENCY

(a) Percent of Decisions in Correct Agreement

No. Analysts Correct	Non-Grain	>50% Grain	>80% Grain	Wheat	Oats	Barley	Summer Crop	Pasture & Grass
0 of 3	1.7	21.5	19.9	7.8	61.4	15.3	0.3	0.0
1 of 3	4.6	13.9	11.2	9.6	10.0	21.2	3.3	0.6
2 of 3	10.7	22.3	20.5	16.6	6.9	34.7	12.2	3.8
3 of 3	83.1	42.4	48.4	66.1	13.6	28.8	84.1	95.6

(b) Percent of Decisions in Correct Agreement as a Random Event (Analyst Accuracies as Prior Probabilities)

No. Analysts Correct	Non-Grain	>50% Grain	>80% Grain	Wheat	Oats	Barley	Summer Crop	Pasture & Grass
0 of 3	0.2	4.9	2.8	0.5	41.1	3.9	.003	.001
1 of 3	1.1	26.3	20.2	7.9	42.8	25.6	15.3	.05
2 of 3	17.4	45.0	45.4	37.4	24.5	48.6	48.3	4.8
3 of 3	81.4	23.8	31.5	54.2	4.1	21.9	84.2	95.2

TABLE 21. PROBABILITY OF CORRECT LABELING AS A FUNCTION OF AGREEMENT

No. Analysts in Agreement	Non-Grain	>50% Grain	>80% Grain	Wheat	Oats	Barley	Summer Crop	Pasture & Grass
2	69.9	61.6	64.7	63.4	40.8	62.1	78.7	86.4
3	98.0	66.4	70.9	89.4	18.1	65.3	99.6	100.0

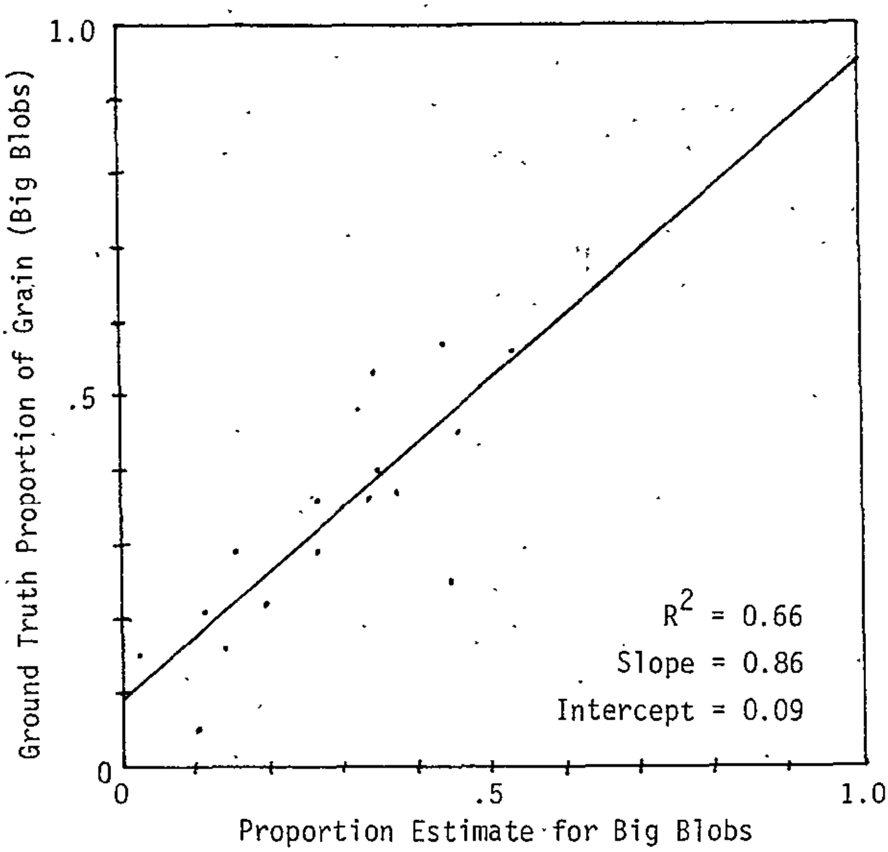


FIGURE 17. PROPORTION ESTIMATES BASED ON ANALYST VOTE LABELS

Analysis of the consistency of proportion estimates between analysts revealed no specific tendencies. Figure 18 illustrates differences in the regression lines of the estimates derived from each analysts labels. On eight occasions the analyst-based estimates were within 0.01 of each other, but these were not in any way related to the degree of accuracy. Interestingly, however, the vote label tended to moderate the level of inaccuracy when analysts' estimates varied substantially from one another. In five segments in which analyst estimates differed by more than 0.13 absolute (an average of 0.165), their accuracy deviated by an average of 0.124 absolute from true, whereas the vote label deviated only 0.070. A reduction in variance is expected in averaging. However, this implies that a vote label results in reduction in the variance of an estimate as well. Accuracy of the vote versus average labels is a function of the confidence one has in the separate labels when a dissenter is present. Table 21 illustrates that in labeling pasture, grass or summer crops, a vote label would have a higher expected accuracy than an average label if there is a dissenter (greater than 66.6%). But in labeling wheat, the dissenter carries more information and an averaging process would be preferable.

4.10 IMPLICATIONS FOR LACIE-IDENTIFIED PROBLEMS

The approach and analysis results presented in the preceding subsections of Section 4 have implications for the possible solution of major labeling problems identified in LACIE (c.f. Section 4.1).

The first and greatest error source identified was "abnormal signatures", that is, temporal sequences of image colors and data that did not match the analyst-interpreter's conception of the proper sequence and pattern for wheat and other small grains. Machine-extracted features and physically-based temporal/spectral stratification offer potential help. While differences in planting date can cause confusing spectral sequences, calculations of spectral crop calendar shift for individual fields or pixels should provide an objective aid for identifying the existence, magnitude, and range of such differences.

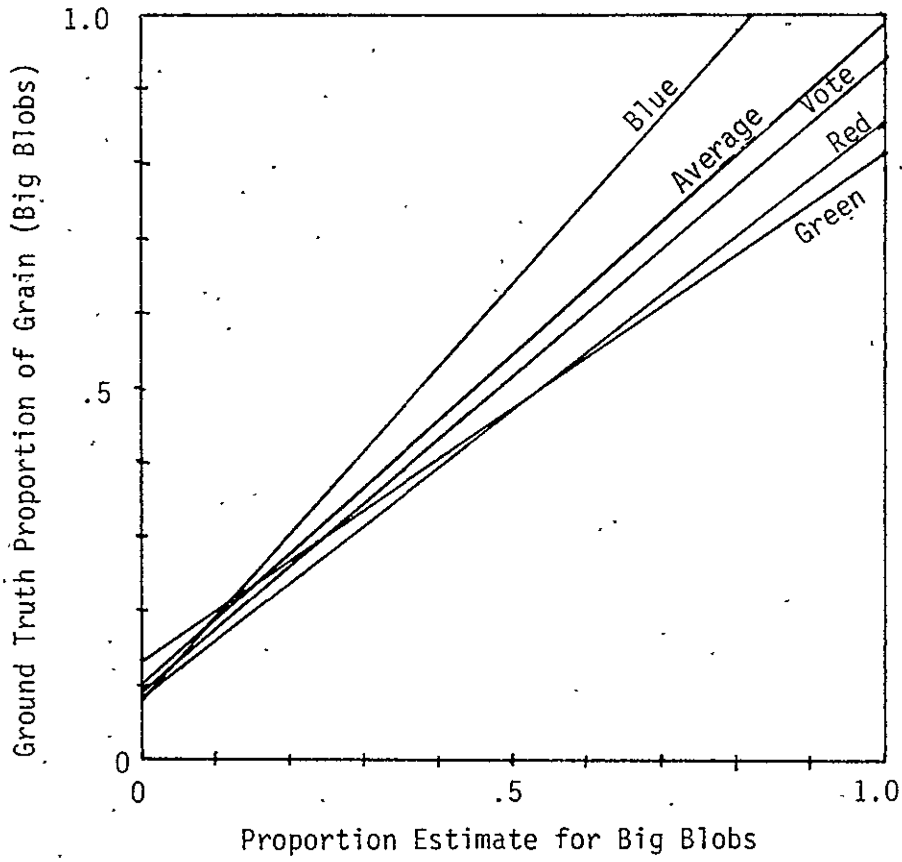


FIGURE 18. REGRESSION LINES FOR PROPORTION ESTIMATES BASED ON ANALYST LABELED BIG BLOBS

Abnormal crop growing conditions, such as drought stress, can cause abnormal sequences and patterns. On a segment-wide basis, spectral indicators such as the Green Index Number and the peak Greenness of a segment-specific profile should help flag those conditions. For individual fields or pixels, profile fits could provide estimates of peak Greenness and green-up and green-down (senescence) rates. In addition to spectral indicators, meteorologically based estimates of crop calendar and moisture stress should be useful for segment characterization and perhaps eventually for field characterization.

Abnormal color patterns in many instances may be artifacts of the color products utilized and the observation conditions. Section 4.6 described an improved approach to the generation of image products that should stabilize the meanings of colors and improve interpretation, consistency and accuracy. Finally, "abnormal signatures" may just be a reflection of a lack of inherent separability in Landsat data. If so, machine approaches will not be of any avail.

The second identified source of labeling errors was associated with boundary and edge pixels. The BLOB algorithm, followed by the stripping operation, allows one to automatically detect and eliminate or otherwise handle such pixels as labeling targets. The consequences of various options for the area estimation process are discussed later in Section 5.5.

The third error source identified was missing acquisitions. Certain combinations must be present to adequately represent the local phenology of the crops of interest. Machine processing cannot create those missing acquisitions, but the temporal-spectral profile fitting technology should help identify the relationships between the crop phenology and the available acquisitions and then provide objective criteria for eliminating segments from further processing if adequate acquisitions are not present.

The final error source was a miscellaneous category that included inconsistent labeling patterns and clerical errors. Machine functions could be defined to do consistency checking and thereby minimize this type of error.

5
MACHINE PROCESSING COMPONENTS FOR AREA ESTIMATION

The components of a stratified area estimation technology were presented in Section 3. The objective of this section is to present developments in the components that create an environment for effective labeling and for the efficient use of labels in overall area estimation technology. Specifically, data normalization (Section 5.2), stratification (Section 5.3), sampling (Section 5.4), and estimation (Section 5.5) are discussed.

A baseline stratified area estimation system, called Procedure M is utilized as a testbed within which to develop and evaluate advanced component technology. Section 5.1 summarizes the salient features of Procedure M to provide a context for the following sections.

5.1 PROCEDURE M

Procedure M [3] embodies a Stratified Area Estimation (SAE) concept. The procedure was developed at ERIM for a multicrop application, based on Procedure B [9] which was developed earlier for application to the estimation of area for a single crop.

LACIE's Procedure 1 [1] is a two-stratum SAE procedure. The statistical framework utilized in Procedure M is in the same multistratum vein as the Procedure 1A approaches currently under development at JSC [33], with the distinction that Procedure M performs a further stratification into large fields and small fields by an automatic technique. In current implementations, only the large fields are used in producing estimates; for them Procedure M is unbiased with respect to the source of labels but, as is typical of existing SAE procedures, errors introduced in labeling are propagated through to the final estimates. A bias is also introduced if the large fields do not appropriately represent the distribution of crops in the small fields and a slight

PAGE 118 INTENTIONALLY BLANK

bias was noted in a major test of the procedure [3]. Sampling strategy is discussed in Section 5.4.

In addition to a field-size stratification, what uniquely sets Procedure M apart is an emphasis on physical understanding related to making area estimates of agricultural classes using remotely sensed data from a space-borne platform. The utilization of field-like targets as the principle sampling unit is an example of this. Another is the emphasis placed on understanding in a formal way the sensor's response to agricultural scenes in order to achieve a standardized spectral domain that is independent of atmosphere, solar geometry and sensor characteristics that interfere with the interpretability of the data in agronomic terms. A spring wheat configuration of the procedure utilizes automatic technology, described in Section 4, to estimate spring wheat in the presence of other spring small grains. This technology is derived from an agronomic understanding of the sensor's response to crop classes and phenologies in the presence of varying soil and climatic conditions.

Table 22 outlines some of the features of Procedure M, and Table 23 describes two specific configurations of Procedure M, as well as indicating areas under development that are discussed in the following sections and elsewhere in this report.

5.2 DATA NORMALIZATION

Landsat data values are measures of radiances in the field of view of the sensor. Efforts conducted in modeling the response of the multispectral scanner [34,35] have illustrated how radiance is affected by atmospheric, geometric and sensor conditions, in addition to the surface phenomenon being viewed. Ideally, a standardized signal response to a given reflectance at the ground is desirable; other variability is just noise, relative to the information need for crop identification. Indeed, even some canopy-related variations can be viewed as noise.

TABLE 22. FEATURES OF PROCEDURE M

- Multiclass

  Stratified areal estimates of any number of classes (e.g., crops) can be produced.

- Multitemporal

  Any number of Landsat acquisitions can be utilized.

- Multisegment

  Segment samples can be grouped in order to reduce the requirement for labeling.

- Modular Implementation

  Procedure components are interchangeable; as components are improved, they are simply inserted in place of existing ones.

- Statistically Stable

  The bias and variance of the estimates are determinable and consistent results are produced to the precision of the mechanism by which the final sampling target is identified (labeled).

- Physically based

  Procedure components take advantage of physical understanding of the sensor and atmosphere, and application-oriented understanding of the resource being inventoried.

TABLE 23. PROCEDURE M COMPONENT TECHNOLOGIES

Component	Multicrop Configuration	Additions for Spring Wheat Configuration	Advanced Technologies Discussed in This Report
Data Preparation	LACIE 5x6 Segments SCREEN $XSTAR_{SV}$		L3 L2 Conversion (Section 5.2)
Feature Extraction	TASCAP BLOB	Trajectory Shift Soil, Stress Indicators	Crop Calendar Estimation SUPERB (Section 4.6)
Stratification	Field Size BCLUSTER		Tolerance Block Physically-based Stratification (Section 5.3)
Sampling	Proportional to Size		Neyman Sampling Label Error Effects (Section 5.3)
Attribute Assignment	Blob Targets Analyst Interpretation	AI for Grains Machine for Spring Wheat	AI Labeling Expmt (Section 4.9) Refined Labeler (Sec 4.8)
Aggregation and Estimation	Weighted Sum		Small Blob Bias Reduction Nearest Neighbor Label Extensions (Sec. 5.5)

The purpose of normalization, then, is to establish a representation of the data that can be interpreted in a consistent manner. Normalization of data in four situations is particularly crucial.

- Sensor-to-Sensor: To utilize similar feature extraction and interpretation technology, independent of the particular measurement system version.
- Segment-to-Segment: To enable the extension of statistical information related to a given crop from one site to another without influence of effects external to the object class of interest.
- Time-to-Time (within a segment): To consistently relate temporal-spectral trajectories of a class (developed from a number of acquisitions) to underlying agronomic phenomena.
- Field-to-Field (within a crop); To investigate the variability of a crop class that is due solely to natural responses to conditions and not due to external effects like the atmosphere.

5.2.1 CURRENT NORMALIZATION PROCEDURES

The Procedure M testbed currently utilizes an integrated set of normalization techniques to reduce variations due to effects external to the canopy itself. The procedures include:

- Data Screening: To remove clouds, shadows, garbled data, and to flag water [36].
- Sun Angle Correction: To standardize the Landsat viewing condition to a fixed solar geometry [30].
- Haze Correction: To standardize the atmospheric conditions at time of acquisition [37].

- Sensor Calibration: To standardize the signal response independent of space-borne multispectral sensor (Landsat 1, 2, and 3) [38, and Section 5.2.2].

5.2.2 LANDSAT 3 TO LANDSAT 2 CALIBRATION

A multiplicative and additive transformation which alters Landsat 3 data to simulate Landsat 2 data was developed this year and is presented in this section. Appendix H describes in detail the procedure employed to develop the transformation. The objective in deriving such a transformation is to be able to directly apply techniques developed for the normalization and interpretation of Landsat 2 data to Landsat 3 data.

The transformation is of the form

$$x_i^{(2)} = a_i\, x_i^{(3)} + b_i \qquad (7)$$

with $x_i^{(2)}$ and $x_i^{(3)}$ representing the Landsat 2 and Landsat 3 signals in channel i.

The vectors $\bar{a}$ and $\bar{b}$ are estimated to be:

$$\bar{a} = \begin{pmatrix} 1.1371 \\ 1.1725 \\ 1.2470 \\ 1.1260 \end{pmatrix} \text{ and } \bar{b} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \text{ where the order is } \begin{pmatrix} \text{MSS4} \\ \text{MSS5} \\ \text{MSS6} \\ \text{MSS7} \end{pmatrix}$$

These coefficients, which indicate an attenuation in each band's signal range, differ by only 2 to 6% from NASA Goddard Space Flight Center prelaunch estimates except for Band 6 which differs by 18%. In addition to the objective statistical procedures used to develop these coefficients, their accuracy was also evaluated subjectively by examining Tasseled-Cap transformed values and the results of applying SCREEN and XSTAR algorithms. The Tasseled-Cap components compared favorably in their orientation and magnitude to Landsat 2 components

of the same segments measured within a nine-day interval. SCREEN and XSTAR, which are based on Landsat 2 calibration, were found to perform reasonably on the corrected Landsat 3 data. Also, Greenness time trajectories composed of both Landsat 2 and 3 acquisitions for particular grain fields were found to be more closely Sigmoidal in shape than their uncalibrated counterparts.

5.3 STRATIFICATION

5.3.1 BENEFIT/COST OF STRATIFICATION

Stratification is a well-known statistical tool to enable efficient sampling of a population. Among several reasons listed for stratifying, Cochran writes [39]:

> Stratification may produce a gain in precision in the estimates of characteristics of the whole population. It may be possible to divide a heterogeneous population into subpopulations, each of which is internally homogeneous...If each stratum is homogeneous, in that measurements vary little from one unit to another, a precise estimate of any stratum mean can be obtained from a small sample in that stratum.

It is for this reason that stratification of Landsat spectral variables is sought.

Efficiency in stratified areal estimates can be realized either by a reduction in the number of labels (measurements) needed for a given variance in the estimate, or by a reduction in variance for a given number of labels.

Richardson [40] suggests a measure of cost/benefit of stratification for a two-class situation (details discussed in Section 6.1.2)* called the fixed-sample reduction of variance factor, RV.

$$RV = \frac{\sum_i \left(\frac{n_i}{n}\right)^2 \frac{p_i(1-p_i)}{a_i} \left(\frac{b_i-a_i}{b_i-1}\right)}{\frac{p(1-p)}{a}\left(\frac{b-a}{b-1}\right)} \tag{8}$$

where n_i is the number of pixels in stratum i,

p_i is the proportion of wheat in stratum i,

a_i is the number of samples allocated to i,

b_i is the size of stratum,

and n, p, a, b are the corresponding segment-level statistics.

This measure assumes sampling proportional to size, equal-sized samples, and a hypergeometric model, and can be simply viewed as the ratio of the variance of stratified random sampling to unstratified random sampling. Therefore, a small RV indicates an effective stratification with strata of high purity.

When the size of strata are very large compared to the sample size, the finite correction terms approach unity and (as Cochran indicates) the RV never increases, and will usually decrease, when any stratum is broken up into smaller strata. However, when the sample size is not negligible compared to the strata size, as is the case in Procedure M, one can conceivably incur a cost in stratification. This is possible if the stratification selected does not provide internally homogeneous strata where they could exist, thereby eliminating certain optimal sample combinations from the selection process. Nevertheless, the average cost over a collection of stratifications will be less than the cost of unstratified sampling [40] as long as the stratification

*Alternative or supplemental information-theoretic measures are also discussed in Section 6.1.2.

is based on features that are actually correlated with the crop-types of interest.

A cost can be incurred, as well, if samples are required to produce the stratification. A simple measure of relative cost for such a procedure can be expressed as

$$\text{Relative Cost} \approx \frac{C_m m + C_n \sum_i n_i}{C_{n_o} n_o} \qquad (9)$$

where n_i is the number of samples allocated to stratum i,

m is the number of samples used to produce the stratification,

n_o is the number of samples required in a simple random sample,

C is the actual cost of the respective sample, and

n_o is chosen to give the same variance in the simple random sample as is achieved using m and n_i in the supervised stratified sample.

In supervised approaches like Procedure M, and potentially, Procedure 1A, m=0 and, in general, $n_i < n_o$ for a given variance. Procedure 1 on the other hand, requires 40 Type 1 dot samples (m=40) to generate initial strata.

Though potentially beneficial, it is critical to select stratification approaches that reduce cost of sampling both in terms of the resultant variance for fixed sample size and in terms of the number of samples required for fixed variance.

5.3.2 STRATIFICATIONS EMPLOYED IN PROCEDURE M

This section will describe and motivate three stratification strategies employed in Procedure M.

1. Stratification of Landsat Pixels into Quasi-Fields: An unsupervised spatial/spectral clustering algorithm called BLOB is utilized to group spectrally homogeneous and spatially contiguous (or nearly so) pixels into quasifields called blobs. This stratification (a) is utilized as a data reduction mechanism--22,000 pixels can be represented by 1000 or fewer blobs, (b) provides sampling units that are more nearly independent than the pixel units, and (c) provides more optimal labeling targets than dots. Section 4.6.3 and Appendix D describes SUPERB, an algorithm to replace BLOB, that insures spatial contiguity of members of a given spatial/spectral cluster. The use of quasi-fields as sampling units is discussed further in Section 5.4.

2. Stratification Based on Quasi-Field Size: Quasi-fields (or blobs) are divided into two categories: Those with at least one interior pixel (large blobs) and those without (small blobs). An interior pixel is one whose four strong neighbors (above, below, left and right) lie in the same quasi-field. The stratum of small blobs is not sampled in current implementations. This eventually leads to a slight potential bias in producing a proportion estimate for the segment, as is discussed in Section 5.5. The small blobs are segregated for two reasons: (1) it is expected that small blobs can be less accurately labeled by analyst interpreters and, hence, any associated label carries less information; and (2) this stratum would capture mixture, misregistered or other unidentifiable blobs that would lessen the homogeneity of spectral strata next described. Testing (Section 4.9) verified that analysts label blobs having only a few interior pixels with lesser accuracy than larger ones. While Statement 2, above, is true, it was found that a number of the small blobs were truly small fields and should be represented in the sampled stratum; this also is discussed in Section 5.5.

3. Stratification of Large Blobs Into Spectral Strata: An unsupervised clustering technique, called BCLUST, utilizes a simple

spectral distance measure and iteratively clusters the means of the large blobs into a fixed (specifiable) number of spectral strata. It is to these strata that samples are directed. Resultant strata have been shown to have fixed sample RV's averaging about 0.5 when acquisitions have been available throughout the growing season. Though encouraging compared to reported Procedure 1 variance reduction estimates, it does not yet compare to the purity levels (RV=0.1) measured for large blob samples that comprise that field-size stratum.

Two approaches have been examined to evaluate the potential of alternative unsupervised clustering techniques to produce strata approaching the RV=0.1 level of purity. The two sections that follow present a method based on statistical criteria and a promising technique based on physical criteria.

5.3.3 STRATIFICATION BASED ON STATISTICAL MODELS

Most methods currently employed in spectral stratification are based on statistical models that seek homogeneous distributions of Landsat spectral multitemporal data. Lennington [41] documents a test wherein three techniques CLASSY, AMOEBA, ISOCLAS, are evaluated. Though the homogeneity of the resultant strata produced by each technique was not compared, it was found that CLASSY produced fewer strata at levels of purity adequate enough to enable more efficient sampling, especially when coupled with sophisticated sample allocation techniques.

To complement the above study, two tolerance block techniques (see Appendix I or [40] for details) and a clustering technique for spectral stratification were evaluated with respect to the estimation of winter wheat acreage in 12 LACIE segments in Kansas in the Procedure M environment. The techniques are (1) to accept tolerance blocks as clusters, (2) to use channel means of tolerance blocks as fixed seeds for spectral clustering, and (3) to conduct unsupervised spectral clustering (BCLUST).

Of the two tolerance block techniques, the seeded clustering tested significantly better as measured by the 100-sample reduction of variance factor. Blocks as clusters produced more evenly sized clusters, which enables efficient sampling, but this partial advantage was more than offset by the lesser spectral homogeneity achieved.

When the tolerance-block-seeded clustering was compared with the unsupervised clustering method BCLUST, there was no significant difference. So in our experiment, the better of the two tolerance block stratification techniques did not show any improvement over previous methods.

A difference exceeding 0.3 remains between the 100-sample RV scores achieved by our two best stratification methods (about 0.5) and what is theoretically attainable, i.e., the score of 0.15 for quasi-field interiors.

The optimal number of strata for a sample of size 100 was not found to be 100 or anything close to it, but rather, 40 for BCLUST and the block-seeded algorithm and 48 for the blocks-themselves algorithm. The reason the optimal numbers weren't higher is that variances produced by the finite sampling create a cost of stratification that can only be made up by purity of strata. In our experiment, 96 fine strata were not enough purer than 40 coarser strata to defray the finite sampling variance cost of the additional strata.

In pursuit of these main conclusions, some subsidiary conclusions were reached.

1. Tolerance block clusters were more uniformly sized than BCLUST clusters, enabling sample allocation to be accomplished more efficiently. However, this advantage did not result in better overall variance reduction.

2. Channels in the first biowindow do help the clustering as applied to winter wheat estimation. The reduction of

variance score for BCLUST averaged 0.048 better when these channels were included.

3. The best channel subsets for generating tolerance blocks contain Brightness and Greenness from the second biowindow.

The tolerance block study could be carried a little further by investigating the use of tolerance block means as seeds and allowing the updating of means and/or cluster creation and/or iteration of clustering. But the payoff from this effort is likely to be small when we compare the distant goal of relatively pure clusters with the modest scores of the clustering methods tested.

A more promising approach based on statistical models would be to redefine features more closely correlated to crop type and test the clustering of these new features using the criterion of the 100-sample reduction of variance factor. The Tasseled Cap features used in the experiment have the virtue of universal applicability. Their use implies only that different materials and crops are localized in separate neighborhoods in spectral space. The relative poorness of the stratification performance indicates the need of features better tailored to the decision problem being considered. Such features could be so specialized that they depend on the crops to be recognized, the confusion crops, the climate, and the prevalent varieties and agricultural practices. If better features are found, there could be a greater reward for dividing the feature space into more homogeneous strata.

The search for features is made in the hope of closing the gap between the RV of 0.5 found for the strata and the RV of 0.15 measuring the purity of the quasi-fields. The possible existance of confusion crops inherently inseparable from wheat could define a higher bound than 0.15 for achievable separability. It may be possible to measure this bound directly, possibly on the basis of a count of identical pairs of data vectors arising from wheat and non-wheat fields, and to chart its

value as a function of the acquisitions available. Such a study would give useful feedback in the search for features and provide a warning when multispectral estimation alone is insufficient.

In addition to the assumption of adequate spectral separability among crop classes, statistically based stratification techniques assume that the underlying statistical distributions are positively correlated to actual crop classes. Certainly, the presence of mixed pixels and misregistered or edge pixels, compound the problem since such pixels may appear statistically identical to crop classes not therein contained.

An alternate approach to statistically based stratification procedures is discussed in the following section.

5.3.4 STRATIFICATION BASED ON PHYSICAL CHARACTERIZATION OF DATA

Rather than employing statistical techniques to determine homogeneous spectral strata, underlying physical characteristics and structure of the data themselves may provide a mechanism for defining temporal-spectral strata homogeneous relative to crop classes.

One example is the Delta Function Stratifier (DFS) [17] developed at the University of California, Berkeley. DFS defines two spectral zones in any Landsat acquisition. These zones are separated by a 'soil' threshold line drawn at the Band 7 over Band 5 ratio equal to 0.55. The pattern of an observation from acquisition to acquisition is used to select strata characteristic of particular crops. We have developed a more detailed temporal-spectral stratification procedure which is described in this section.

Physically based procedures may rely on an assumption that the objective criteria for stratifying or 'zoning' data are standard from data set to data set. The ratio used in DFS in the vicinity of the 'soil' line is less sensitive to external effects than the individual bands, since many of those effects are in the direction of the soil line.

The Procedure M environment proves to be an ideal one in which to apply and evaluate physically based temporal-spectral stratification approaches, since extensive efforts are made to normalize data to a standard reference as described in Section 5.2.

There is an intuitive appeal to stratification strategies which are based on physical understanding. DFS represents an initial venture into this domain. However, more sophisticated procedures based on more detailed features are possible. A study of this potential has been initiated and shows promise, not only in providing a low cost means to define reasonably homogeneous strata, but also an environment to employ more sophisticated sampling strategies to increase the overall efficiency of the system, as will be presented in Section 5.4.

Figure 19 illustrates four zones established for purposes of stratification of spectral data. The zones are separated by dashed lines superimposed, in the Brightness-Greenness plane, on thresholds used for screening each acquisition. These four zones were based on an understanding of crop spectral phenology and established empirically by observation of critical spectral stages in the development of small grain fields in 13 North Dakota blind sites. Temporal sequences of these zones then can define temporal-spectral strata for area estimation purposes.

The first zone is associated with early phenological stages of a grain field, such as bare soil and limited vegetation cover, as well as later stages of late senescence and harvest. The second zone is a transition zone through which the grain spectra pass as they 'green-up', i.e., develop toward maximum ground cover and heading; they also transition through this zone as they senesce and the grain ripens. The third zone includes the green arm with high vegetation cover and vigorous growth. The fourth zone was defined to include any fields that transitioned at high levels of Brightness.

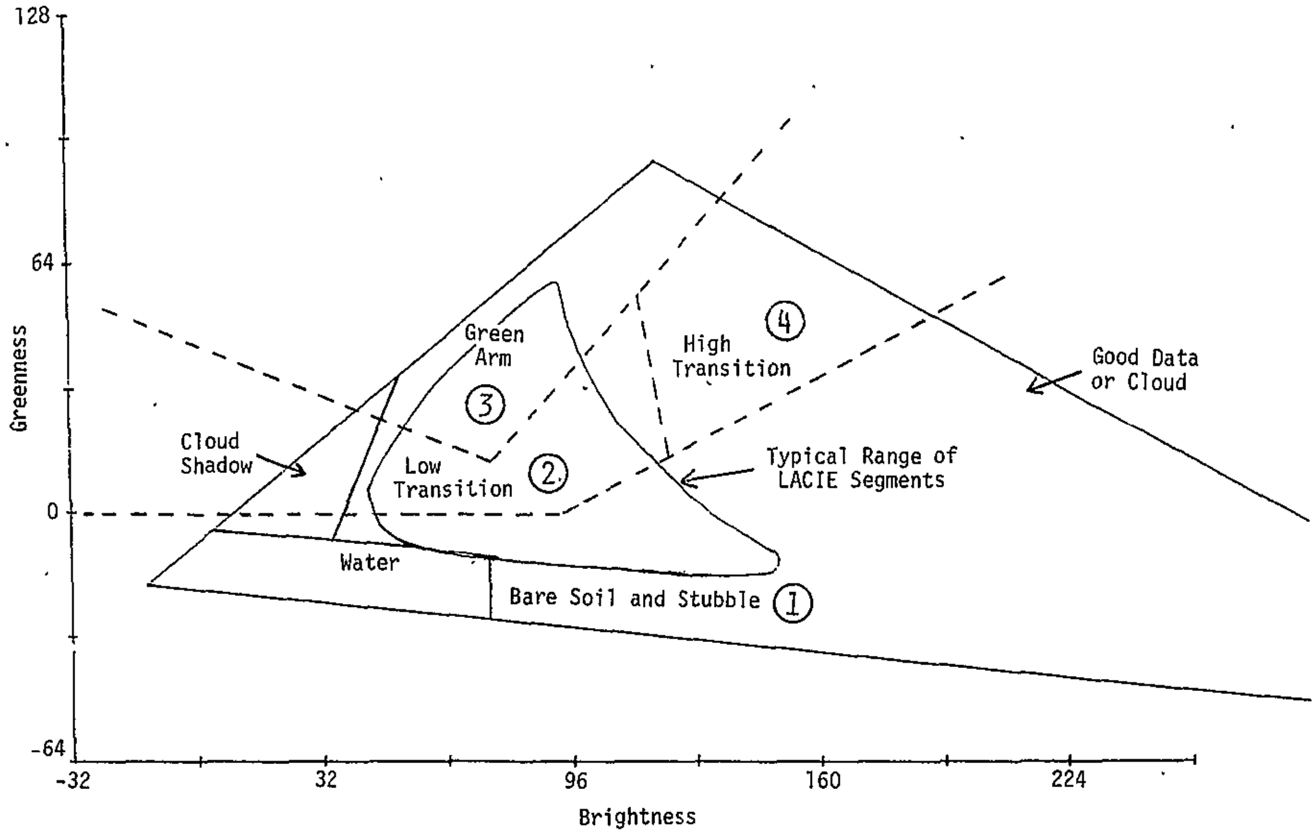


FIGURE 19. PRINCIPLE ZONES OF CROP DEVELOPMENT FOR THE TEMPORAL/SPECTRAL STRATIFICATION

All zones are restricted to the envelop of screening thresholds that encompass good data. Within this constraint, the zones are given by the following relationships:

Let $\bar{x}_i$ be a normalized Landsat field mean or signal value for Acquisition i.

Let g_x, b_x be the Greenness and Brightness values of $\bar{x}_i$, each with a +32 count offset from the values presented in Figure 19. Then within the triangle of good data:

$$
\left.
\begin{array}{l}
\bar{x}_i \subset \text{Zone 1 if} \\
\qquad g_x \leq 32.40 \\
\quad \text{or} \\
\qquad g_x \leq 0.382\, b_x - 7.00 \\
\\
\bar{x}_i \subset \text{Zone 3 if} \\
\qquad g_x \geq -0.326\, b_x + 67.00 \\
\quad \text{and} \\
\qquad g_x \geq b_x - 47.58 \\
\\
\bar{x}_i \subset \text{Zone 2 if} \\
\qquad \bar{x}_i \not\subset \text{Zone 1 or Zone 3} \\
\quad \text{and} \\
\qquad g_x \leq -0.872\, b_x + 149.49
\end{array}
\right\} \quad (10)
$$

$$
\begin{array}{l}
\bar{x}_i \subset \text{Zone 4 if} \\
\qquad \bar{x}_i \not\subset \text{Zone 1 or Zone 2 or Zone 3}
\end{array}
$$

Multitemporal strata are defined by determining the sequence of zones through which a spectral trajectory passes. For example, if acquisitions were collected at critical times, the sequence 1,2,3,2,1 would represent the stratum associated with fields that together traverse a normal development cycle from bare soil, to peak green, senescence and harvest.

Spectral stratifications based on these zones were examined in light of their resultant homogeneity with respect to spring small grains.

In three North Dakota segments examined, 1392, 1473, and 1653, the stratification resulted in strata that were remarkably pure with respect to grain and nongrain distributions. Tables 24, 25, and 26 illustrate resultant strata using three acquisitions. The meaning of these tables is discussed in the paragraphs that follow.

The stratum purity levels achieved are noted in three stages, one for each of three acquisitions used. Potentially four strata will be formed in the first acquisition. Each of these is potentially divided into four more, and so on. As many as $4^3 = 64$ strata are possible here; however, many of these multitemporal strata were not populated. It will be the concern of future investigations to establish primary strata and to evaluate techniques to combine less populated strata. Ideally acquisitions should be grouped into biowindows before stratification so that strata are consistent from segment to segment. However, this initial examination was segment specific.

Comparing the proportion of grain in each stratum to the overall segment proportion, each subsequent stage achieves purer strata. The purity levels achieved are shown in three stages, indicating how each stage contributes additionally to purifying the final strata.

TABLE 24. STRATIFICATION TREE FOR SEGMENT 1473

Grain Proportion: 0.58*; State: North Dakota

Julian Date

197	207	224
Zone 1 Grain Proportion 0.50 Size (% of Pixels) 14.07%	Zone 1 0.0 5.8%	0.0 5.8%
	Zone 2 1.0 5.96%	1.0 5.96 %
	Zone 3 0.37 3.01%	0.96 1.18%
		0.0 1.85%
	Zone 4	

197	207	224
Zone 2 0.65 11.30%	Zone 1 .92 5.52	0.92 5.52%
	Zone 2 0.57 3.66%	1.0 1.82%
		0.14 1.85%
	Zone 3 0.11 2.12%	1.0 .24%
		0.0 1.88%
	Zone 4	

197	207	224
Zone 3 0.54 60.6%	Zone 4 0.0 0.25%	0.0 0.25%
	Zone 2 .93 22.16%	1.0 18.97%
		0.51 3.05%
		0.0 0.13%
	Zone 3 0.47 21.98%	0.94 2.97%
		0.81 9.44%
		0.0 8.47%
		0.0 1.10%
	Zone 4 0.11 16.23%	1.0 1.85%
		0.0 0.18%
		0.0 4.21%
		0.0 9.99%

197	207	224
Zone 4 0.74 13.99%	Zone 1 1.0 8.93%	1.0 8.93%
	Zone 2 1.0 0.15%	1.0 0.15%
	Zone 3	
	Zone 4 0.26 4.91%	1.0 1.27%
		0.0 0.60%
		0.0 0.40%
		0.0 2.63%

*Only large blob strata considered.

NOTES: (1) A perfect stratification has grain proportions of 1.0 or 0.0 in each cell.
(2) The numbers in each box represent, respectively, the grain proportion (fraction) and segment area (%) contained in that stratum.

TABLE 25. STRATIFICATION TREE FOR SEGMENT 1392

Grain Proportion: 0.38*; State: North Dakota

Julian Date

154	190	208
Zone 1 0.61 22.68%	Zone 1 0.0 2.48%	Zone 1 0.0 1.68%
		Zone 2 0.0 0.21%
		Zone 3 0.0 0.6%
		Zone 4
	Zone 2 0.0 1.43%	Zone 1
		Zone 2 0.0 0.13%
		Zone 3 0.0 0.3%
		Zone 4 0.0 1.01%
	Zone 3 0.73 18.77%	Zone 1 1.0 2.04%
		Zone 2 .79 4.77%
		Zone 3 .79 8.16%
		Zone 4 .40 3.79%
	Zone 4	Zone 1
		Zone 2
		Zone 3
		Zone 4
Zone 2 0.54 34.77%	Zone 1 .04 8.08%	Zone 1 0.01 6.70%
		Zone 2 0.37 0.70%
		Zone 3 0.0 0.68%
		Zone 4
	Zone 2 0.24 3.20%	Zone 1
		Zone 2 0.27 2.83%
		Zone 3 0.0 0.19%
		Zone 4 0.0 0.19%
	Zone 3 .75 23.42%	Zone 1 1.0 1.87%
		Zone 2 0.76 9.56%
		Zone 3 0.93 7.32%
		Zone 4 0.33 4.66%
	Zone 4 1.00 0.07%	Zone 1
		Zone 2 1.0 0.07%
		Zone 3
		Zone 4
Zone 3 0.12 42.55%	Zone 1 0.0 5.50%	Zone 1 0.0 3.69%
		Zone 2 0.0 1.80%
		Zone 3
		Zone 4
	Zone 2 0.005 23.13%	Zone 1 0.0 0.75%
		Zone 2 0.0 20.83%
		Zone 3 .13 0.85%
		Zone 4 0.0 0.71%
	Zone 3 0.37 13.76%	Zone 1 0.57 1.26%
		Zone 2 .45 6.67%
		Zone 3 0.19 4.34%
		Zone 4 0.32 1.50%
	Zone 4 1.0 0.15%	Zone 1 1.0 0.15%
		Zone 2
		Zone 3
		Zone 4
Zone 4	Zone 1	
	Zone 2	
	Zone 3	
	Zone 4	

*Only large blob strata considered.

TABLE 26. STRATIFICATION TREE FOR SEGMENT 1653

Grain Proportion: 0.15*; State: North Dakota

Julian Date

155	191	208
Zone 1	Zone 1	0.0 1.2%
	0.00	0.0 0.2%
	1.4%	
Grain Proportion	Zone 2	0.0 0.3%
0.60	0.35	0.63 0.8%
	1.5%	0.00 0.4%
Size (% of Pixels)	Zone 3	0.94 2.1%
6.9%	0.90	0.86 2.0%
	4.1%	
	Zone 4	

155	191	208
Zone 2	Zone 1	0.0 1.3%
	0.0	0.0 0.31%
	1.5%	
	Zone 2	0.0 2.7%
0.15	0.04	0.03 45.0%
	48.5%	0.65 0.79%
61.43%		
	Zone 3	
	0.65	0.71 8.7%
	11.2%	0.46 2.6%
	Zone 4	
	0.0	
	0.09%	

155	191	208
Zone 3	Zone 1	
	0.0	0.0 0.13%
	0.13%	
	Zone 2	0.0 0.15%
0.06	0.04	0.04 20.9%
	21.87%	0.1277 0.79
31.4%		
	Zone 3	0.0 0.11%
	0.11	0.07 7.05%
	8.9%	0.29 1.69
		0.0 0.11%
	Zone 4	
	0.0	0.0 0.06%
	0.45%	0.0 0.06%

155	191	208
Zone 4	Zone 1	
	Zone 2	
0.0	0.0	
	0.06%	
0.18%		0.0 0.06%
	Zone 3	
	0.0	
	0.11%	
		0.00 0.11%
	Zone 4	

*Only large blob strata considered.

This trend is summarized on Table 27 using a purity measure defined as:

$$P = \frac{\Sigma_i w_i \max (p_{Gi}, p_{NGi})}{\sum_i w_i} \tag{11}$$

where

- P is the purity factor
- w_i is the stratum size
- p_{G_i}, p_{NGi} is the proportion of grain or non grain in stratum i

The first stage of stratification results in a purity level about equal to that of the whole segment; each subsequent stratification results in a marked improvement in overall purity, each new stratum being more homogeneous with respect to grains.

It remains to evaluate the physically based approach to temporal-spectral stratification in comparison to statistical methods. Indeed, a combined method may prove useful. Certain strata can be refined as a function of time of year. Combining of strata may be necessary to reduce the overall number possible, with no loss in homogeneity. Strategies to stratify multiple segments with different acquisition histories are required. Finally, modification of physically based stratification strategies using physical implications of collateral conditions can eventually be expected to yield improved results over those achievable using purely static definitions.

Initial signs for the potential of physically based stratification are promising. The next section on sampling discusses an important benefit of such stratification related to employing efficient sampling strategies.

TABLE 27. STRATIFICATION PURITY* LEVELS

Segment	Proportion of Grain**	Average Purity Level of Stratification		
		Stage 1 (1 acq)	Stage 2 (2 acqs)	Stage 3 (3 acqs)
1392	.38	70.13%	84.13%	85.79%
1437	.58	57.48%	84.55%	99.13%
1653	.15	86.05%	91.15%	92.82%

$$^{*}\text{Purity} = \frac{\sum_i w_i \max(p_{G_i}, p_{nG_i})}{\sum_i w_i}$$

** Proportion of large blob stratum.

5.4 STRATEGIES FOR SAMPLING OF FIELDS

ERIM has long held the view that near-term improvements in estimating the proportion crops of interest within a segment can be accomplished using stratification based on spectral variables and the sampling of fields. Procedure M, in particular, employs stratification using spectral variables of field-like forms. Section 5.3 introduced the concept of strata which are comparable, or static, from segment to segment. The main advantages of such a stratification strategy are:

1. The grain proportion within a spectral-temporal stratum should show some stability from segment to segment. One might view the proportion of grain within a stratum as having a distribution depending on the stratum.

2. The labeling errors within a stratum could also be similar from segment to segment. One would expect each stratum would have its own confusion crops, for example.

Thus one would hope that information concerning the stratum distributions of grain and labeling errors could be ascertained by experiment, and that such information could be used in optimizing sample allocation and possibly in bias correction of the sample proportion of grain labels. Appendix N gives some approaches on how this information could be used. This section will provide a summary.

5.4.1 CURRENT PROCEDURE M STRATEGY

The stratification used by Procedure M is a clustering procedure using Brightness-Greenness variables subject to a constraint on the number of clusters. The stratum sample size is proportional to the number of pixels in the stratum. This is a slight generalization on the Cochran allocation in which the samples are allocated proportional

to the population of the stratum. The segment proportion estimate is the weighted average of the strata proportion estimates where the weights are the relative sizes of the strata.

The stratum proportion estimate is obtained by using the Midzuno sampling procedure. This procedure chooses the samples in such a way that the field sizes are taken into account. This is done in a fashion which at first glance is unintuitive. The first field from a stratum is chosen with probability proportional to its area. The remaining are chosen wich equal probability. The net effect of this sampling procedure is that the probability of obtaining any one particular sample is proportional to the area of the fields in that sample.

We now give an example to illustrate this point. Suppose that there are three fields: F_1, F_2, and F_3. Suppose further that two of these fields are to sampled and the sizes are 5, 3, and 2 pixels, respectively. Figure 20 gives a tree diagram in which the first field is chosen proportional to number of pixels and the second is chosen from the remaining two fields with equal probability. In this simple example, there are only three possible samples: (F_1,F_2), (F_1,F_3), and (F_2,F_3). We now note from the figure that:

$$P(F_1 \text{ and } F_2) = P(1^{st}\ F_1 \text{ then } F_2) + P(1^{st}\ F_2 \text{ then } F_1)$$
$$= \frac{5}{20} + \frac{3}{20} = \frac{8}{20}$$

$$P(F_1 \text{ and } F_3) = P(1^{st}\ F_1 \text{ then } F_3) + P(1^{st}\ F_3 \text{ then } F_1)$$
$$= \frac{5}{20} + \frac{2}{20} = \frac{7}{20}$$

$$P(F_2 \text{ and } F_3) = P(1^{st}\ F_2 \text{ then } F_3) + P(1^{st}\ F_3 \text{ then } F_2)$$
$$= \frac{3}{20} + \frac{2}{20} = \frac{5}{20}$$

The possible samples, their probabilities, and the number of pixels contained in each sample follows:

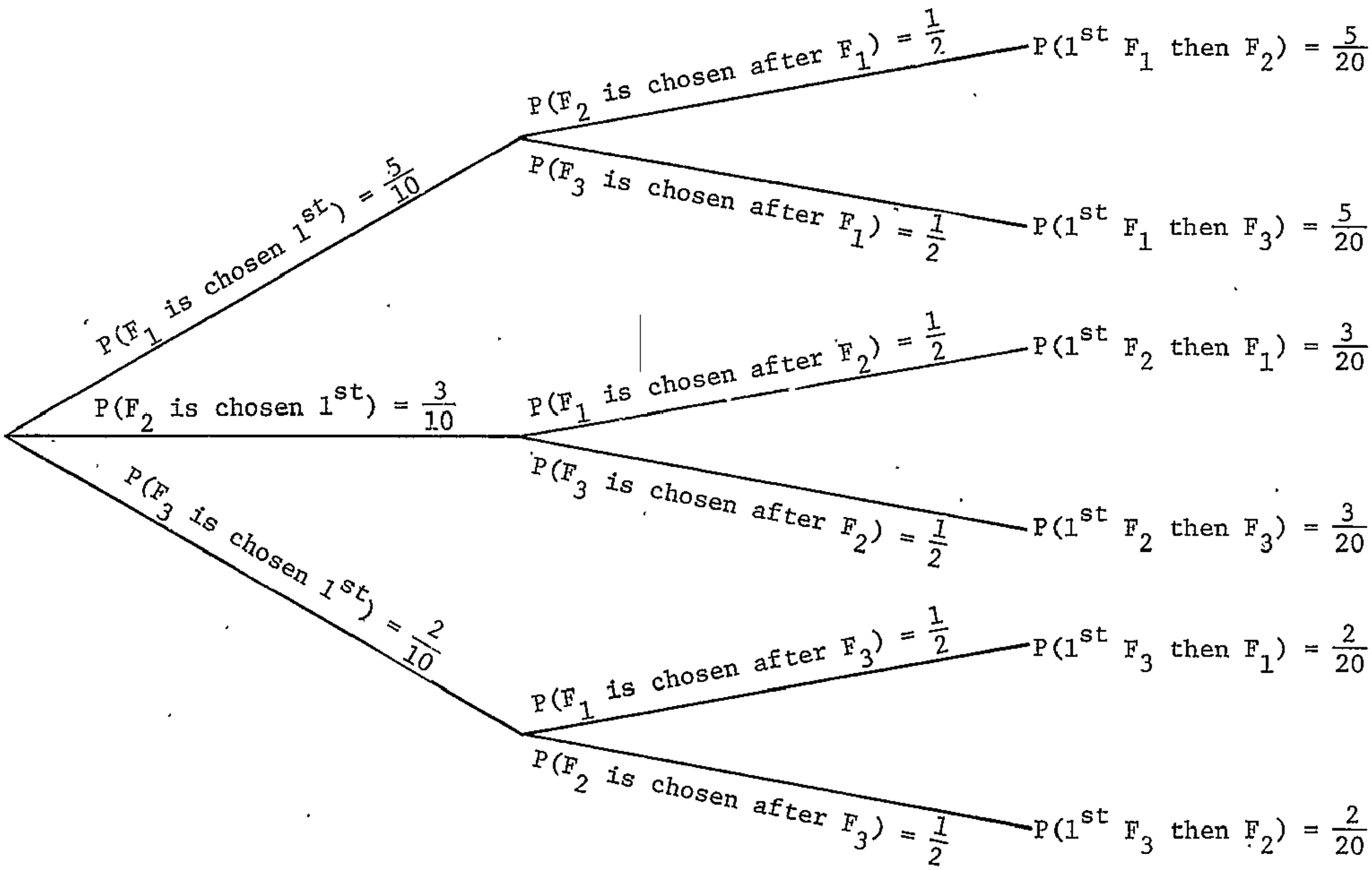


FIGURE 20. MIDZUNO SAMPLING OF 2 FIELDS FROM A POPULATION OF 3 FIELDS OF SIZES 2, 3, and 5 PIXELS

Sample	P(Sample)	Area of Fields in Sample
F_1, F_2	$\frac{8}{20}$	8 Pixels
F_1, F_2	$\frac{7}{20}$	7 Pixels
F_2, F_3	$\frac{5}{20}$	5 Pixels

Note that the probability of obtaining any one sample is proportional to the number of pixels contained in that sample.

5.4.2 ALLOCATION OF SAMPLES TO STRATA

In this section, we present some approaches to the problem of allocation of a fixed sample size to the strata in order to minimize expected sample variance, as alternatives to the current allocation scheme in Procedure M. This approach depends upon having a stratification that remains sufficiently stable from segment to segment so knowledge about each stratum can be accumulated. We assume that the total sample size is a fixed constraint, say n. We have an estimator $\hat{P}$ which is a linear combination of the strata proportion estimators $\hat{P}_i$, namely

$$\hat{P} = \sum_i \left(\frac{N_i}{N}\right)\hat{P}_i \qquad (12)$$

where N_i is the number of pixels in stratum i, and $N_. = \Sigma N_i$. The variance of $\hat{P}$ is

$$V(\hat{P}) = \sum_i \left(\frac{N_i^{\,2}}{N_.^{\,2}}\right) V(\hat{P}_i) \qquad (13)$$

If the sampling cost per unit is the same in all strata, then theorem 5.7 of Cochran states:

> In stratified random sampling $V(\hat{P})$ is minimized for a fixed total size of n if

$$n_i = n \; \frac{N_i\sqrt{V_i(\hat{P}_i)}}{\sum_j N_j \; V_j(\hat{P}_i)} \qquad (14)$$

This allocation is sometimes called "Neyman Allocation" after Neyman (1934) who gave one of the first proofs of its optimal properties among the class of unbiased. Neyman allocation has the obvious drawback that one usually does not know the true variances of the strata. There are several ways to adjust for this lack of knowledge:

a) Allocate proportional to stratum size, N_i/N. This sometimes is called Cochran allocation. You will recall that this allocation is currently used in Procedure M. Cochran gives in his theorum 5.8:

   If terms in $1/N_i$ are ignored relative to unity,

   $$V_{optimum} \leq V_{proportional} \leq V_{random}$$

   where the optimum allocation for fixed n is $n_i \propto N_i \sqrt{v_i}$.

b) Allocate proportional to $(N_i/N)\ B_i$ when $V(\hat{P}_i) \leq B_i$. This allocation is a minimax allocation in the sense that the allocation is made according to the worst possible stratum variance. This allocation is discussed in Appendix N.

c) Allocate proportional to expected variance with respect to prior distribution, τ_i, on each stratum, namely:

   $$n_i = \frac{n\ E_i(V)}{\sum_j E_j(V)} \qquad (15)$$

   where E_i is the expected value under the probability measure τ_i.

5.4.3 STRATEGIES WHICH USE KNOWLEDGE OF LABELING ERRORS

If the distributions of the strata labeling errors are estimable from segment to segment then information on labeling errors can be used to allocate samples to strata. Since labeling error can introduce variance to an estimate, it is the duty of the allocation procedure to

be cognizant of this source of variance. We consider sample allocation in the presence of labeling error in this section.

Suppose we have a stratum which as P proportion of grain targets of equal size. Also suppose that there are labeling errors; that is, grain targets are labeled as grain with probability 'α' while nongrain targets are labeled as nongrain with probability 'δ'. This concept is displayed in Figure 21. When ground truth is not considered then the labels have errors associated with them. The number of grain labels obtained from a sample of size n is a binomial $(n, P\alpha + (1-P)(1-\delta))$ random variable (Figure 22.) The labeling introduces bias and variance into the estimation process. Some possible methods of obtaining, updating, and using joint strata priors on (P,α,δ) are given in Appendix N.

The mean squared error of P is the sum of the variance and squared bias of $\hat{P}$, that is

$$MSE(\hat{P}) = V(\hat{P}) + b^2(\hat{P}). \tag{16}$$

Since the sample size has no effect on b(P) only the strata variances are used to allocate samples to strata. However if there were good priors for some of the strata in which labeling errors were high, then it might be the case that the expected MSE is reduced when no samples are allocated to those strata and the priors for P_i are used in place of $\hat{P}_i$. This, of course, requires that the strata are stable with respect to (α,δ,P).

If within strata the distributions of α and δ have small variances then the bias of the stratum proportion estimate might be reduced by using the estimate

$$\hat{P}_s = \frac{\bar{P}_s - (1 - \bar{\delta}_s)}{\bar{\alpha}_s + \bar{\delta}_s - 1} \tag{17}$$

where $\bar{P}_s$ is the sample proportion of strata,

$\bar{\alpha}_s$ is the mean of the prior distribution of strata for α_s, and

$\bar{\delta}_s$ is the mean of the prior distribution of strata for δ_s.

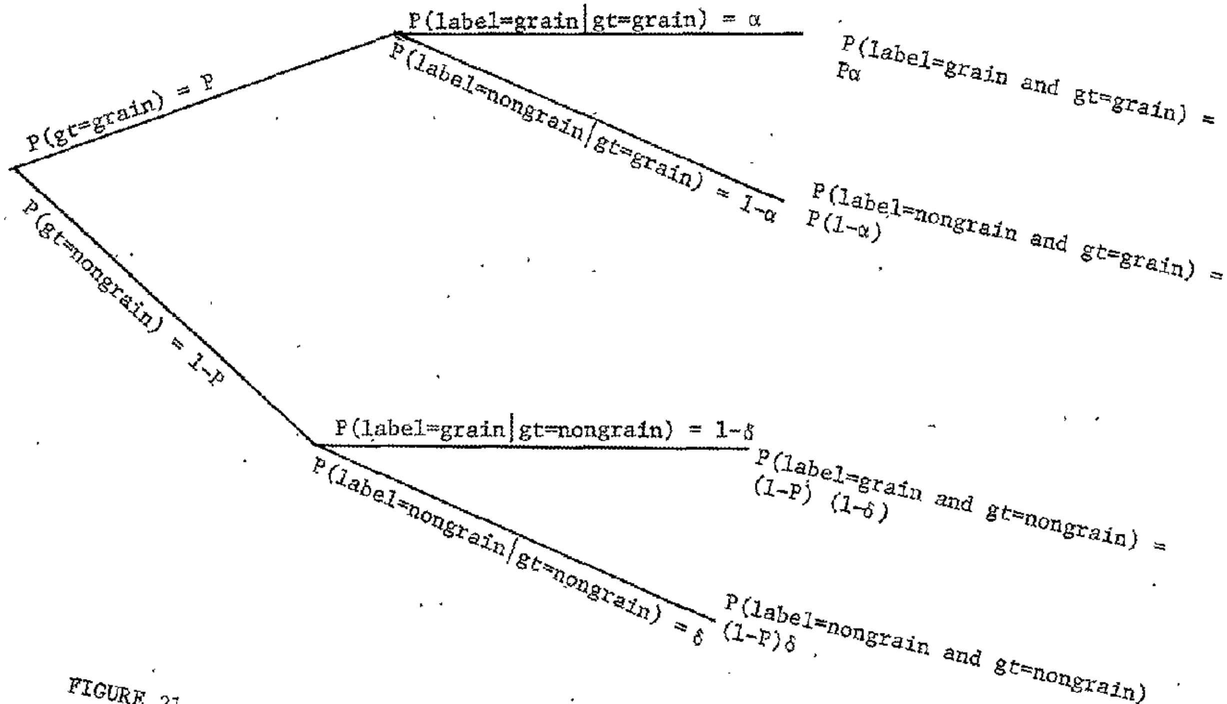


FIGURE 21. FOUR POSSIBLE OUTCOMES IN SAMPLING WITH LABELING ERRORS

$$P(\text{label=grain}) = P\alpha + (1-P)(1-\delta)$$

$$P(\text{label=nongrain}) = P(1-\alpha) + (1-P)\delta$$

FIGURE 22. TWO POSSIBLE OUTCOMES IN SAMPLING WITH LABELING ERRORS WHEN ONLY THE LABELS ARE KNOWN

5.5 ESTIMATION

Using Procedure M the estimate of the proportion of crops in each spectral stratum is the weighted aggregation of the labeled samples in the stratum:

$$p_i^{(j)} = \frac{\sum_k n_{ik} \, p_{ik}^{(j)}}{\sum_k n_{ik}} \tag{18}$$

where $p_i^{(j)}$ is the estimated proportion of crop j in stratum i

$p_{ik}^{(j)}$ is the proportion of crop j in the k^{th} sample of stratum i

n_{ik} is the size of the k^{th} sample.

The estimate associated with the entire segment is:

$$p^{(j)} = \frac{\sum_k m_k \, p_k^{(j)}}{\sum_k m_k} \tag{19}$$

where $p^{(j)}$ is the estimated proportion of crop j in the segment

m_k is the size of each stratum

The estimate $p^{(j)}$ is unbiased with respect to those strata sampled. In current implementations of Procedure M, a potential bias is introduced due to not sampling the stratum of small blobs and can be estimated as:

$$b^{(j)} = \frac{N-M}{N} \left(P_s^{(j)} - P_u^{(j)}\right) \tag{20}$$

where $b^{(j)}$ is the bias in estimating crop j

$P_s^{(j)}$ is the proportion of j in the large blob stratum from which samples are drawn

$P_u^{(j)}$ is the proportion of j in the unsampled stratum

N is the total number of pixels

M is the total number of pixels in the sampled stratum

The sampling variance of the procedure has been empirically estimated in Procedure M. An analytic expression for the variance (which is the variance associated with Midzuno sampling) is discussed in Section 6.

In this section we will examine approaches considered to reduce the bias due to not sampling the stratum of small blobs, and an alternate estimation strategy based on a non-parametric classification strategy that potentially reduces sampling variance.

5.5.1 CONTROL OF PROCEDURE M SAMPLING BIAS

Equation 20 indicates that a potential bias is introduced in Procedure M that is due to not sampling the stratum of small blobs. This bias is related to the relative size of this stratum and the difference in crop composition between the sampled and unsampled strata. Empirical tests in the Great Plains have shown that the relative bias is typically on the order of 5%, with the relative size of the small blob strata at 20% of the segment. Several approaches for reducing this bias are suggested.

(1) Sample Little Blobs: Upon observation of blob boundary maps produced for the AI experiment (Section 3) many little blobs are bypassed as potential labeling targets, even though they seem reasonable to label and clearly visible on imagery. That is, small fields are represented by little blobs that are not boundary blobs, misregistered blobs, or miscellaneous blobs.

It is possible that our definition of labeling targets as the set of big blobs is unduly restrictive. One feasible approach is as follows:

(a) Stratify the scene into big and little blobs as usual.

(b) Stratify the little blobs further to produce a stratum of little blobs with more than n pixels.

(c) Sample the big blob strata and the strata of (b), above proportional to size. Direct the samples for the big blob strata to spectral strata as usual. The little blob strata should be treated separately, either sampling randomly or after some form of spectral stratification.

(d) Let n be the samples allocated to the little blob strata. This strata may still include blobs that are very difficult to identify. Provide the analyst m x n blobs (m = 2?) in some random order. The analyst should label at least n of these blobs.

(e) Aggregate the results as usual.

This approach will reduce the size of the unsampled strata thus reducing the sampling bias. However, the interaction of field size and analyst labeling error may result in a labeling bias. The trade off involved has yet to be evaluated.

(2) Contextual Stratification: Cochran writes: "Sampling problems may differ markedly in different parts of the population. With human populations, people living in institutions are often placed in a different stratum than people living in ordinary homes because a different approach to the sampling is appropriate for the two situations" [39]. The following suggests an approach that stratifies the quasi-fields according to contextual information, then treating each stratum according to techniques best suited for that stratum.

The strategy consists of several steps. The context of each small blob (the identity and spectral characteristics of neighboring blobs) is determined. Based on whether the small blob arises due to a multitemporal misregistered boundary, a spectral anomaly (e.g., pond, group of trees), a small field, or other cause, appropriate action is taken, so that the area of the small blobs can correctly be assigned to the crop categories. Appropriate action might consist of a classification technique that takes advantage of knowing the nature of spectral contamination or mixing (and possibly contamination due to spatial misregistration) that is present.

The specific approach follows:

(a) Examine the effect of adjusting BLOB parameters on the prevalence of small blobs and on the resulting bias. The results of running BLOB with a variety of parameter settings can be examined with respect to purity, blob size, and number of blobs with only edge pixels. If these results indicate the likelihood that the bias and overall system performance will be held to acceptable levels, then tests of the procedure will be carried out using the new settings. Alternatively, other approaches must be sought.

(b) Study the nature of the small blobs, and the physical situations that lead to their existence. Since it is anticipated that different effects may give rise to small blobs -- a local unusual spectral phenomenon, boundaries especially if spatial misregistration is present, strip fields, and valid small fields -- it is important to understand the relative occurrence of these potential causes. It must be determined whether these causes can be distinguished using only the scene data. The question of whether BLOB boundaries are accurately placed must also be addressed. The problems involved with spatial registration of time periods must be considered in this context.

(c) After certain insight is gained regarding the small blobs, relationships can be sought between the observed bias and various features. These features may include relative field size distributions of respective crops (or of spectral strata), segment crop proportions, and other ancillary information. It has been shown, for example, that bias is a strong function of wheat proportion [3].

To attain confidence in relationships found, physical causes should be sought. This confidence based on physical insight is required if the relationships are to be trusted for any other data set than the ones initially used for development. Testing should be done by applying the relationships to sets of segments not involved in deriving the relationships.

(d) At this time, it is not known whether Step c will produce a solution that will be considered reliable in general. Therefore, a more involved approach may be required:

Since the small blobs will not be sampled and analyst-labeled, they must be classified or otherwise dealt with. However, several difficulties with classification must be overcome. Four general steps are as follows.

First, training of the classifier must be carried out. In the past, selection of representative training samples was a perennial problem, but with the advent of Procedure M, samples selected and labeled will provide representative training material.

Second, raw classification is usually a biased procedure. A Monte-Carlo or other procedure to measure classifier bias should be implemented, and the result used to correct the classification bias.

Third, the context of each small blob should be examined. This context includes the spectral composition of the neighboring blobs, perhaps their identity, but certainly the conditions which lead to the formation of the small blob and the degree of presence of spatial misregistration. This context can be used to determine how to process the small blob, and where appropriate, can provide information not previously available to a classifier (e.g., boundary location, adjacent field signals).

The final step depends on the nature of the small blob. If it is a misregistered boundary of two larger fields, the area would be assigned to them. If it is some other anomalous situation, appropriate action would be taken. If on the other hand it represents a bonafide small field, then the situation deserves paragraphs of its own.

A small field may be covered by a group of pixels none of which are pure as illustrated in Figure 23. However, the signal statistics of the neighboring blob and the position of the actual boundary are presumably known for each pixel, so that a maximum likelihood procedure can be used to produce the observed pixel. The aggregate of several of these estimates that comprise a small blob may be sufficiently pure to allow classification to take place.

This is a complex problem that may require accurate knowledge of misregistration, accurate boundary location, and other techniques. Its effectiveness is probably reduced somewhat in areas where small blobs adjoin small blobs. But it does address the small fields problem in greater detail than has been usual.

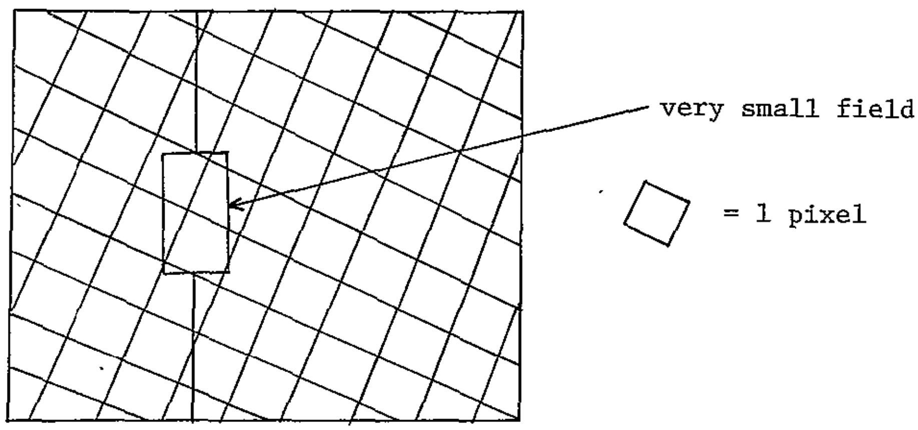


FIGURE 23. SMALL FIELD REPRESENTED BY ONLY MIXTURE PIXELS

This contextual approach is highly sophisticated, suggesting a merging of classification schemes into a stratified area estimation system. It suffers in that little is known about modeling the bias and variance characteristics of sophisticated classifiers and in the complications of developing appropriate criteria for contextual stratification.

5.5.2 A NONPARAMETRIC APPROACH TO LABEL EXTENSION

Work by S. Wheeler at IBM illustrates an approach to area estimation using a simple linear discriminant and a cross validation technique [5] to monitor its bias and variance performance. This section discusses an approach utilizing a nonparametric classifier with implications to the contextual stratification approach identified above (details provided in Appendix M).

Appendix M provides a framework for crop proportion estimation based on a nonparametric nearest neighbor decision rule. With nearest neighbor assignments, unlabeled units are 'classified' by assigning the label of the nearest labeled sample. Advantages of nearest neighbor assignment are that no assumptions are made regarding the distributions of crop classes, and the performance of the classifier can be estimated from the training samples.

This label extension approach has ramifications on contextual stratification proposed in Section 5.5.1. Labels derived from samples in the big blob stratum can be extended to the stratum of unsampled small fields. The requirement imposed upon this approach is that the resultant 'labels' associated with the small blobs are more accurate than those resulting from direct labeling. Short of complete classification, nearest neighbor extension of labels can be restricted by a specified distance to increase the probability of a correct classification. The resultant set of samples, composed of the original training sample and the classified sample, can then be used to estimate proportion of crops present. The nearest neighbor label extension approach is under continued investigation.

6
PERFORMANCE EVALUATION

Section 5 discusses a baseline Procedure M technology for stratified estimation of crop acreage with Landsat, as well as some advances proposed for the technology. This section presents analytic and empirical evaluations of the baseline procedure. Section 6.1 presents analytical modeling considerations and Section 6.2 summarizes the status of an empirical test of the procedure conducted using 1978 segments from the Northern Great Plains.

6.1 THEORETICAL CONSIDERATIONS

As noted in Section 2, performance models have three major roles in crop inventory systems. Generically these are:

1. Self-assessment (of an operational system).
2. Acceptance test design.
3. Acceptance test evaluation.

All of these functions can take place at various levels of a crop inventory system, from the overall system, to major components, to various subcomponents. As a natural part of the process of developing an area estimation component we have been constructing error models to go with each component or sub-component. Specifically, we have constructed error models covering the area estimation component up to the segment level.

It is desirable to characterize component performance as a function of a variety of collateral conditions; specifically the collateral conditions ought to consist of the variables used as stratifiers at the next higher level of aggregation in the system. So far, we have not conducted sufficient testing to be able to establish reliable associations with collateral variables.

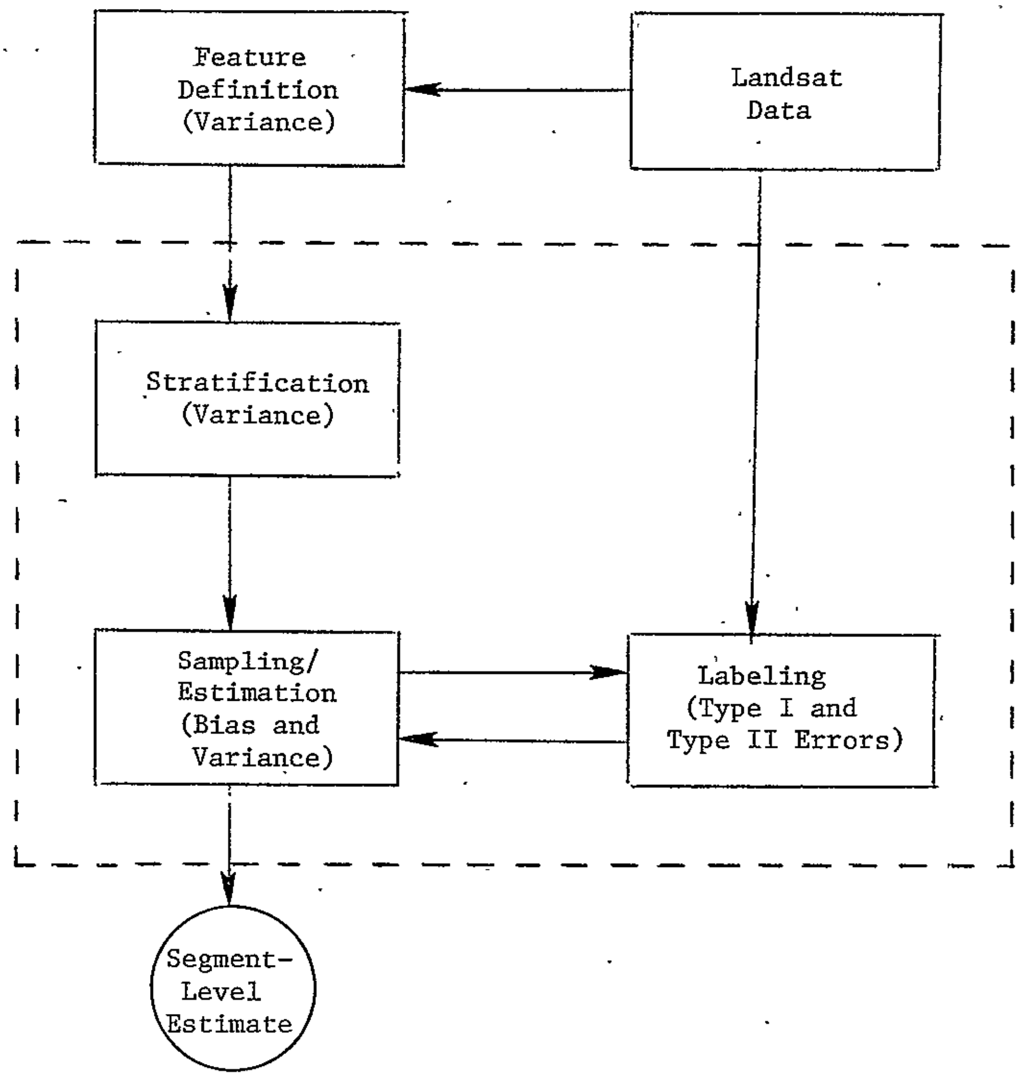
PAGE 158 INTENTIONALLY BLANK

The performance measures that ought to be used in error modeling are not definitely established. To date we have used the bias and the covariance of proportion estimates as the measure of performance, and also "purity", and reduction of variance factors. An alternative concept is the entropy of estimates, discussed under "Information Theoretic Performance Measures", in Section 6.2 and Appendix K. Varieties of information theoretic measures have been applied to Procedure M in order to obtain some insight into these measures.

In Section 6.1.1 following we describe the overall structure of performance models developed so far for Procedure M, up to the segment level. Section 6.1.2 discusses performance measures including reduction of variance and information theoretic measures.

6.1.1 ERROR MODELING (SEGMENT LEVEL, PROCEDURE M)

Figure 24 is a flow diagram of Procedure M as applied at the segment level. The input is multitemporal Landsat data. The initial steps of processing include screening, instrument calibration, haze correction, spectral feature extraction (Tasseled-Cap linear features) and spatial feature extraction (the production of quasi-fields through the use of the BLOB algorithm). This is all indicated by the box called Feature Extraction. Following feature extraction, the blobs are stratified on the basis of their spectral properties and size characteristics. This is indicated by the box called stratification. Sample blobs are drawn from each stratum according to a specific sampling scheme and the selected blobs are presented to a labeler, either an analyst or a computer algorithm. The samples are labeled and then aggregated to produce strata estimates and finally a segment-level estimate.



The dashed lines surround the components for which error models have been formulated.

FIGURE 24. STAGES IN PROCEDURE M WHICH ARE THE SUBJECT OF AN ERROR MODEL

The types of error which may arise and propagate through the system are also shown in the figure. The sampling/estimation box shows bias and variance. For every sampling scheme there exists some estimation scheme which produces an unbiased estimate, but not every estimation scheme is unbiased, and some estimation formulas produce higher variance than others. The currently used technique (the Midzuno sampling technique) produces unbiased estimates for the strata sampled, assuming accurate labeling.

The process of sampling of course introduces variance. Various sampling techniques (Cochran, Neyman) are discussed elsewhere in this report. A principle assumption is that one can direct sampling proportional to the size of a stratum or proportional to some other property of the stratum. For finite strata, this is of course not exactly possible. The variance expression contains both a finite correction term (hypergeometric distribution) and a correction for integer round-off in the number of samples drawn.

The process of stratification does not produce any bias. A good stratification is one which reduces variance of the overall estimates by increasing the purity of the strata created compared to the original unstratified purity. Hence stratification affects variance and a measure of goodness of stratification is the reduction of variance factor, RV, calculated under the assumption of ideal sampling.

Feature extraction cannot introduce bias; however there is information loss (which we believe to be minimal) in the feature extraction process in Procedure M. This information loss can introduce variance as noted in the figure.

The dotted lines in the figure surround the portions of the system for which error models have been developed or are being studied.

Considering the performance of stratification alone we have defined the "fixed sample reduction of variance" which is developed and presented in the next section and in Appendix J, under the assumption of perfect labeling and ideal proportional sampling. The contribution of labeling error to bias and variance is developed for two basic sets of assumptions about labeling; first that the Type I and Type II labeling error rates are fixed, and second that the error rates themselves are drawn from a random distribution whose statistics are fixed. These expressions are summarized in the next section and are developed in detail in Appendix J.

The Midzuno sampling scheme for sampling with unequal sized samples does not have a simply expressed variance formula, and for most evaluations the simpler models in the following section have been employed. However a variance expression does exist. It is presented in the experimental evaluation of an 18-segment Procedure M test with analysts described in Appendix C.

6.1.2 PERFORMANCE MEASURES

Several measures of performance have been utilized in evaluating Procedure M technology. These measures are utilized in grading not only the overall performance of the procedure, but the contribution of each component. These measures are presented here due to their general applicability to the evaluation of crop acreage inventory technology. The principal measures include a simple purity measure, and variance reduction factor [40] (Section 6.1.2.1) and measures based on information theoretic considerations (Section 6.1.2.2). The next two sections describe advancements made in defining the performance measures. Appendix K describes information theoretic measures in detail.

6.1.2.1 The Fixed Sample Reduction of Variance Factor

Reduction of Variance Factor (RV)

The measure of performance heretofore used [42] to evaluate clustering parameters and methods is the reduction of variance factor

$$RV = \frac{\sum_{\text{all strata } i} n_i p_i (1 - p_i)}{np(1 - p)} \tag{21}$$

where n_i is the number of pixels in stratum i,

p_i is the proportion of wheat in stratum i,

n is the number of pixels in the segment $(n = \Sigma n_i)$,

p is the proportion of wheat in the segment $(p = \Sigma n_i p_i / n)$.

The RV is the ratio of two variances: the variance of the stratified sample estimate divided by the variance of the unstratified sample estimate. It is a number between 0 and 1. A small number is good, indicating that the stratified estimate has a proportionally smaller variance than the unstratified estimate and so the stratification is doing some good. We can verify in Expression 21 that if the strata are either pure wheat or pure other, then either p_i or $1 - p_i$ is 0 and the numerator is 0. If the stratification is worthless, then the p_i's are all the same as p and the factor becomes 1.

RV With Integer Allocations

The RV as a performance measure is unrealistic in two ways. For one thing, it assumes that we are allocating the sample exactly in proportion to the size of the strata. Such an allocation is optimal in the absence of information about the true percent wheat p_i in each stratum. But it is an approximation because the number of quasi-fields sampled from a stratum must be an integer whereas with few exceptions, the exact proportional allocation is not an integer.

The approximation becomes absurd when the number of strata increases beyond the size of the sample. Then strata must be sampled with a probability rather than with certainty and the variance should rise. But the simple expression (21) does not take account of this effect and continues to decrease (get better) as the number of strata increases.

The approximation is not burdensome when we compare results for clustering algorithms producing approximately equal numbers of strata. But when the numbers are unequal, as when we are trying to find the optimal number of clusters for a given algorithm, the comparison is invalid.

So we can define a better performance measure by assuming a realistic sample size, say 100 quasi-fields, and allocating them to strata as best we can, that is, as nearly as possible proportional to size. If some strata are left unallocated, we'll combine them into a wastebasket stratum and sample it. Then,

$$RV = \frac{\sum_{\text{strata } i} \left(\frac{n_i}{n}\right)^2 \frac{p_i(1 - p_i)}{a_i}}{\frac{p(1 - p)}{a}} \tag{22}$$

where n_i is the number of pixels in stratum i,

p_i is the proportion of wheat in stratum i,

a_i is the number of sample quasi-fields allocated to stratum i,

and n, p, a are the corresponding numbers for the segment.

The allocations $\{a_i\}$ are made by a subroutine as follows:

1. Determine the theoretical allocation an_i/n for each stratum i.
2. Round this number to the nearest integer.

3. Collect all the strata with allocation 0 into a wastebasket stratum and allocate sample quasi-fields to it proportional to size, but at least 1. Thus no strata are left out of the sampling.

4. If the integer allocations don't add to a, multiply the fractional allocations by $1 + \varepsilon$ and repeat. ε is chosen by an algorithm that makes the procedure rapidly converge. There are, however, some numerical combinations that prevent convergence, and then we settle for an allocation that doesn't quite add up to a.

The RV with integer allocation (22) is not likely to improve as the number of strata exceeds the sample size because the number of terms being summed in the numerator of (22) remains constant and the wastebasket stratum, in all probability heterogeneous, increases in size.

The Fixed-Sample RV

A second unrealistic assumption in using expression (21) is sampling with replacement. In fact, it is only reasonable to assume sampling without replacement, implying a hypergeometric, rather than a binomial model. The effect on the RV is to multiply numerator and denominator by correction factors as follows:

$$\text{Fixed-Sample RV} = \frac{\sum \left(\frac{n_i}{n}\right)^2 \frac{p_i(1 - p_i)}{a_i} \left(\frac{b_i - a_i}{b_i - 1}\right)}{\frac{p(1 - p)}{a} \left(\frac{b - a}{b - 1}\right)} \tag{23}$$

where n_i is the number of pixels in stratum i,
p_i is the proportion of wheat in stratum i,
a_i is the number of sample quasi-fields allocated to stratum i,
b_i is the number of quasi-fields in stratum i,
and n, p, a, b are the corresponding numbers for the segment.

This is the realistic performance measure that is used for comparing clustering methods. It is still an approximation because it assumes that all sample quasi-fields are the same size.*

An implication of the finite correction factors is that stratification incurs a cost. Let us illustrate by an example. Suppose that we create 100 strata, so evenly divided that we allocate one sample quasi-field to each stratum. The correction factor in the numerator is always 1 and drops out. In the denominator, b, the number of quasi-fields might typically be 400, so the correction factor is 3/4. Now suppose that the stratification completely fails to discriminate, so that p_i is constantly equal to p. Then everything cancels out but the 3/4 and we are left with a reduction of variance factor of 1 1/3! This means the variance of the stratified estimate is 1/3 more than that of the unstratified estimate. Stratification hasn't helped in this case!

This example is extreme because if the stratification were made at random, then just by chance we would expect most p_i's to be different, then p and perhaps some to be close to 0 or 1. So two opposing forces influence stratification: the finite correction factors penalize stratification and discrimination of wheat from non-wheat rewards it. If the stratification is made at random, it has been shown that the two forces would be expected to approximately cancel each other out [43]. This is reasonable; one would expect that a random stratification followed by a random sampling from the strata would be equivalent to a random sample from the whole population.

If there are labeling errors these are propagated through the Procedure M stratified sampling and estimation scheme. Both bias and variance are created by labeling errors. Let α be the average rate of correct classification of wheat as wheat and β be the average rate of misclassification of non-wheat as wheat. Then the bias is

$$E(\hat{p} - p) = \frac{\sum n_i\left(\left[\alpha_i - \beta_i - 1\right]p_i + \beta_i\right)}{\sum n_i} \tag{24}$$

and the variance is

$$E(\hat{p} - E(\hat{p}))^2 = \sum\left(\frac{n_i}{n}\right)^2 \frac{A_i + B_i}{a_i} \tag{25}$$

where:

$$A_i = (\beta_i - \beta_i^2) + (\alpha_i - \beta_i)(1 - \alpha_i - \beta_i)p_i \tag{26}$$

$$B_i = (\alpha_i - \beta_i)^2 pi(1 - p_i)\left(\frac{b_i - a_i}{b_i - 1}\right). \tag{27}$$

This formula is developed from the general expression for propagation of variance; if

$$L = f(w)$$

and

$$\mathrm{Var}(L|w) \text{ is defined}$$

then

$$\mathrm{Var}\ L = E\left(\mathrm{Var}(L|w)\right) + \mathrm{Var}\ E(L|w) \quad .$$

The first of these terms corresponds to the A_i, the second to the B_i in Equation 25.

In general we do not expect that the error rates, $1 - \alpha$ and β, are accurately known, but it may be possible to establish statistics on those error rates, i.e., we may be able to estimate a mean and covariance for the α and β. In this case the bias and variance of Procedure M will be estimated using a further expansion of the general formula for variance. The expression is given in Appendix J, Equations J-18 and J-19. Full development of these expressions is given in Appendix J.

Previously we stated that a system performance model ought to output the covariance of the crop estimates in the multicrop case. If the crop proportion p is regarded as a vector and $\hat{p}$ likewise, then Equation 23 can be interpreted as the $(j,k)^{th}$ term of the covariance of p by replacing $p_i(1 - p_i)$ by $(-p_i^{(j)}p_i^{(k)})$ and replacing $p(1 - p)$ by $(-p^{(j)}p^{(k)})$.

Let π be the multicrop confusion matrix such that π_{kj} is the rate of misclassification of class j as class k. Then Equation 24 can be written as

$$E(\hat{p} - p) = \frac{\sum n_i(\pi_i - 1)p_i}{\sum n_i} \tag{28}$$

and each of the more general cases can be directly expanded to the covariance expression.

6.1.2.2 Information Theoretic Performance Measures

Current performance measures for agricultural inventory applications of remote sensing systems are their accuracy and precision in crop area estimation (i.e., bias and variance) and their probability of correct classification. Intermediate stratification steps are measured by their variance reduction factors.

It was conjectured that the well-developed principles of information theory should be applicable to evaluation and design aspects of remote sensing systems that extract information and estimate crop area and production.

The objective of the effort summarized here and detailed in Appendix K was to determine the validity of this conjecture and, if valid, to develop performance measures to supplement or parallel the current measures. An investigation was conducted, an approach was established, and several performance measures were defined. Only a very limited amount of empirical evaluation was conducted, but the results are encouraging.

The approach taken was to view information extraction systems as being communications channels with scene characteristics at the input and derived or estimated attributes or characteristics at the output. Figure 25 illustrates the concept. Note that the processing system in general is not perfect and introduces noise. This noise can cause information to be lost in transit or can introduce errors in the output. A more detailed consideration of input/output pairings is presented in Appendix K.

Table 28 summarizes the basic information theory concepts upon which the figures of merit are based. Entropy and mutual information are the quantities of interest.

Several figures of merit for stratification procedures are developed and discussed in Appendix K; all have values that range from zero to unity and are normalized values of mutual information. They differ in their normalizing factors. The one that appears most appropriate normalizes by the total entropy, thereby being sensitive to both the number

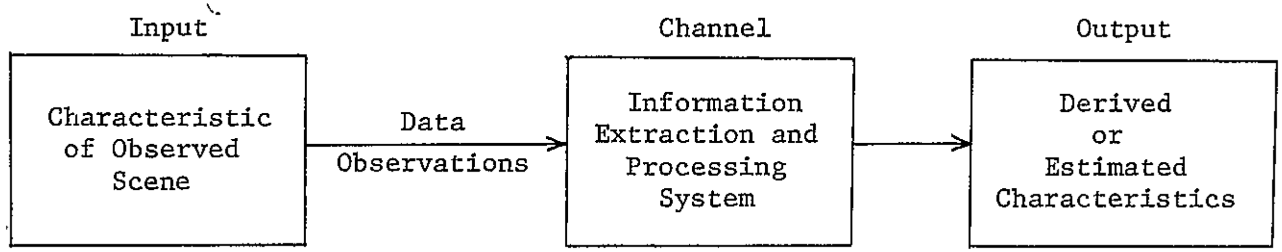


FIGURE 25. ILLUSTRATION OF THE COMMUNICATIONS CHANNEL CONCEPT FOR INFORMATION EXTRACTION SYSTEMS

TABLE 28. BASIC INFORMATION THEORY CONCEPTS

(1) The self information associated with the occurrence of state x_i which occurs with probability $P(x_i)$ is defined to be:

$$I(x_i) = \log \frac{1}{P(x_i)} = -\log P(x_i)$$

(The more rare the event, the greater is the amount of information associated with its occurrence.)

(2) Entropy is the average amount of information associated with repeated observations of a state variable:

$$H(X) = -\sum_{i=1}^{m} P(x_i) \log P(x_i)$$

$$\left(H_{max} = \log m \text{ and occurs when } P(x_i) = \frac{1}{m}\right)$$

(3) Mutual information is the expected average information exchanged from input X to output Y:

$$I(X;Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} P(x_i,y_j) \log \left(\frac{P(x_i|y_j)}{P(x_i)}\right)$$

of output strata and their purity. Performance curves for a two-class input are shown in Figure 26 as a function of number of output strata for three levels of output stratum purity. The penalty paid by impurities of as little as 5% is striking, but also in accord with results using reduction of variance measures (See Section 6.1.2.1). Also, note that a change from 2 to 4 output strata reduces $M_{X,Y}$ about the same as a reduction of purity from 95 to 90%. Incorporation of a cost factor could allow a different weighting function.

Information theoretic performance measures have several potential advantages. A principal one is that they are directly extendable to multiple crop situations. Also, information gain can be quantified in additive units. More analysis is required to evaluate the usefulness of the developed stratification figures of merit relative to, or as supplements to, the variance reduction factor; incorporation of cost functions may also prove desirable. The limited result presented in Appendix K is promising. It is recommended that investigation of the use of information-theoretic concepts be continued and extended to other aspects of information extraction system performance and evaluation, such as classification and sampling.

6.2 EVALUATION OF PROCEDURE M COMPONENTS

The configuration of Procedure M, described in Section 5.1, for a spring small grains application is undergoing evaluation using analyst labels for small grains, machine labels for spring wheat, and ground truth labels as a baseline for comparison. The experiment design is described in Appendix C. The purpose of the evaluation is to demonstrate the performance of Procedure M in the presence of analyst labels, and to establish the relative merit of each component of this stratified area estimation procedure. This section will summarize the status of the evaluation of two components of the Procedure -- BLOB, which defines quasi-field principle sample units, and BCLUSTER, a clustering algorithm used for spectral stratification.

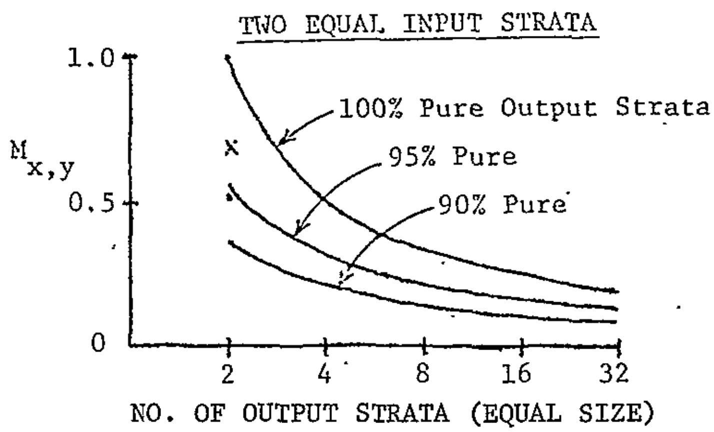


FIGURE 26. INFORMATION THEORETIC PERFORMANCE CURVES FOR A TWO-CLASS STRATIFIER

6.2.1 QUASI-FIELD DEFINITION WITH BLOB

The BLOB algorithm has undergone both the subjective scrutiny of three LACIE-experienced LEC analysts and quantitative evaluation. BLOB [44] is a multitemporal spatial/spectral clustering algorithm that forms field-like patterns of Landsat pixels that are used as the primary sample unit and labeling target in Procedure M. These patterns are referred to as 'quasi-fields' or 'blobs'. Acquisitions used in the algorithm were selected by analysts as to their merit with respect to distinguishing field features. In this experiment, no specific guidelines were developed for this selection, but were left to the discretion of the LEC analysts. A summary of the results of the two evaluations conducted follows.

Analysts were to:

- Visually determine if blobs correspond to actual fields.
- Evaluate if the choice of acquisitions used in BLOB were adequate.
- Indicate particular undesirable artifacts or anomalies of the algorithm that are observed.

Details of analysts' responses are presented in Appendix C. The salient points of their response are:

- Blobs visually correspond to actual fields in most segments.
- A breakdown of this correspondence in certain segments could be atributed to an inappropriate selection of acquisitions; guidelines as to the appropriate selection were suggested but not utilized.
- Certain blobs were disjoint,* i.e., not all pixels contiguous. This was found to be an undesirable feature, especially in labeling.

---

*SUPERB, described in Appendix M, addresses modification to the BLOB algorithm that eliminates this problem.

Quantitative evaluation of BLOB performance is expressed in terms of a simple purity measure. We define purity with respect to a canopy as the percentage of a blob that is comprised of the crop or canopy of interest.

In 17 LACIE TY 78 blind sites in the Northern Great Plains, over 6,000 blobs averaged 93% pure as grain or non-grain. Figure 27 histograms blob purity to illustrate that 80% of the blobs were at least 80% pure with respect to this two-class categorization.

Blobs that were at least 80% pure as spring small grains numbered 1531. A large majority of these (1235, or 81%) were in addition at least 80% pure with respect to a specific spring small grain. In all cases, including cultivated and non-agricultural canopies, a large number of blobs were found to be at least 80% pure and an average of greater than 90% pure with respect to a specific canopy. Table 29 lists purities for a number of ground truth categories.

6.2.2 DEFINITION OF SPECTRAL STRATA WITH BCLUSTER

The BCLUSTER algorithm [9] is a simple unsupervised clustering algorithm that is currently utilized in Procedure M to form spectral strata using means of blobs contained in the big blob strata. BCLUSTER can be controlled to produce any predefined number of strata or 'BCLUSTERS'. This section presents an evaluation of spectral strata produced for 13 North Dakota TY 78 blind sites using three stratification levels -- 20, 40 and 60 strata per segment.

The simple variance reduction and purity factors will be the performance measures discussed. The average simple variance reduction factors resulting from forming 20, 40 and 60 strata were .637, .544, and .483 respectively. Table 30 presents the segment-by-segment breakdown.

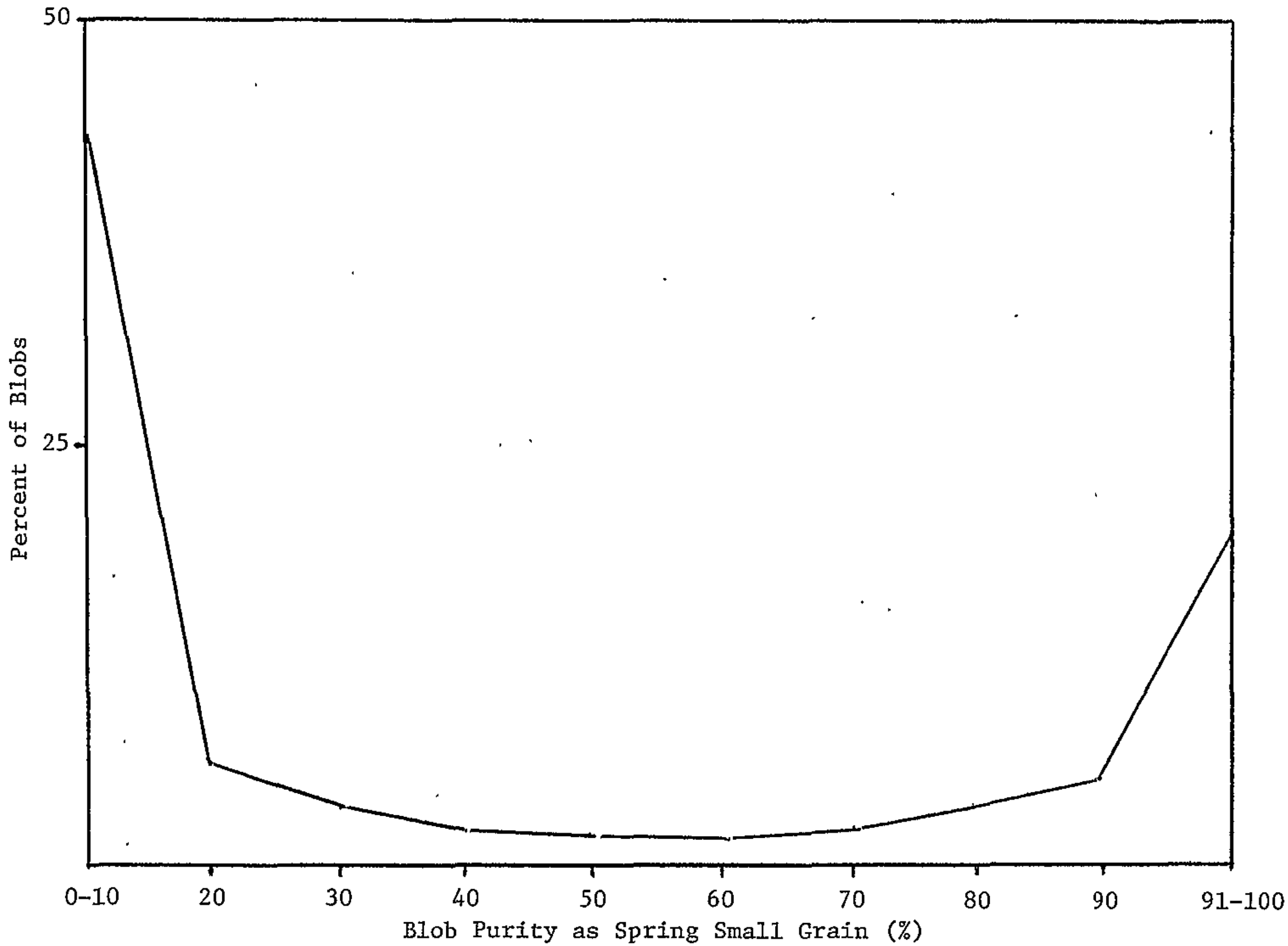


FIGURE 27. BLOB PURITY FOR 17 NORTHERN GREAT PLAINS TY78 SEGMENTS

TABLE 29. BLOB PURITY LEVELS

Ground Truth Category		Average Purity*	No. Blobs
Generic	Specific		
Spring Crops		93.3	1531
	Spring Wheat	91.1	800
	Barley	92.6	154
	Oats	96.4	267
	Rye	93.0	14
Summer Crops		94.4	730
	Alfalfa	91.1	42
	Corn	90.4	137
	Soybeans	94.8	178
	Sunflower	93.7	187
Not Cultivated		N/A	N/A
	Pasture	93.9	393
	Fallow	91.8	537

*Average purity among those blobs at least 80% pure within the generic category specified.

TABLE 30. SIMPLE VARIANCE REDUCTION FACTORS OF 13 TY 78 BLIND SITES BASED ON SPECTRAL CLUSTERING OF BIG BLOB MEANS WITH BCLUSTER

	Number of Strata		
Segment	20	40	60
1392	0.709	0.479	0.422
1457	0.562	0.516	0.441
1461	0.715	0.504	0.482
1467	0.723	0.683	0.621
1473	0.441	0.335	0.315
1602	0.463	0.434	0.371
1612	0.701	0.654	0.627
1619	0.490	0.342	0.292
1636	0.705	0.616	0.469
1650	0.728	0.634	0.616
1653	0.657	0.613	0.500
1656	0.731	0.652	0.541
1920	0.654	0.608	0.577
Average	0.637	0.544	0.483

This result indicates a substantial improvement in efficiency in stratification relative to simple random sampling. Though the result indicates that 60 BCLUSTERS are most efficient, the gain in purity is offset by allocating a fixed sample. For example, the distribution of 100 samples over 40 spectral strata could be accomplished more nearly proportional to size than over 60 strata. The fixed sample RV (Section 6.1.2.1) provides a means to compute an RV relative to a given allocation, but was not needed for the purposes or this analysis.

Figure 28 illustrates the trend toward efficiency as more strata are formed. The purity of BCLUSTERS relative to their grain or non-grain composition is histogrammed. The percentage of strata that are relatively pure non-grain (greater than 80%) remains relatively constant independent of the number of strata targeted. This implies a significant level of separability between certain grains and non-grains. However, the percentage of relatively pure grains shows a dramatic increase from 20 to 40 BCLUSTERS. The implication is that a large percentage of grains and non-grains are spectrally close, and it is not until the finer threshold levels required to produce more strata are utilized that the grain and non-grain separate.

Comparison of Figure 28 to comparable illustrations used in the evaluation of Procedure 1 [10] and P1A [33,41] imply improved stratification with BCLUSTER over ISOCLAS, AMOEBA and CLASSY. It is conjectured, however, that much of the apparent improvement is due to the use of blobs instead of pixels and also to excluding the stratum of little blobs in the stratification, rather than to an improved spectral clustering procedure. It is recommended that comparisons of clustering procedures be based on data that is not contaminated by mixed or misregistered pixels since these pixels can be confused as canopies that do not contribute to the signal's spectral composition.

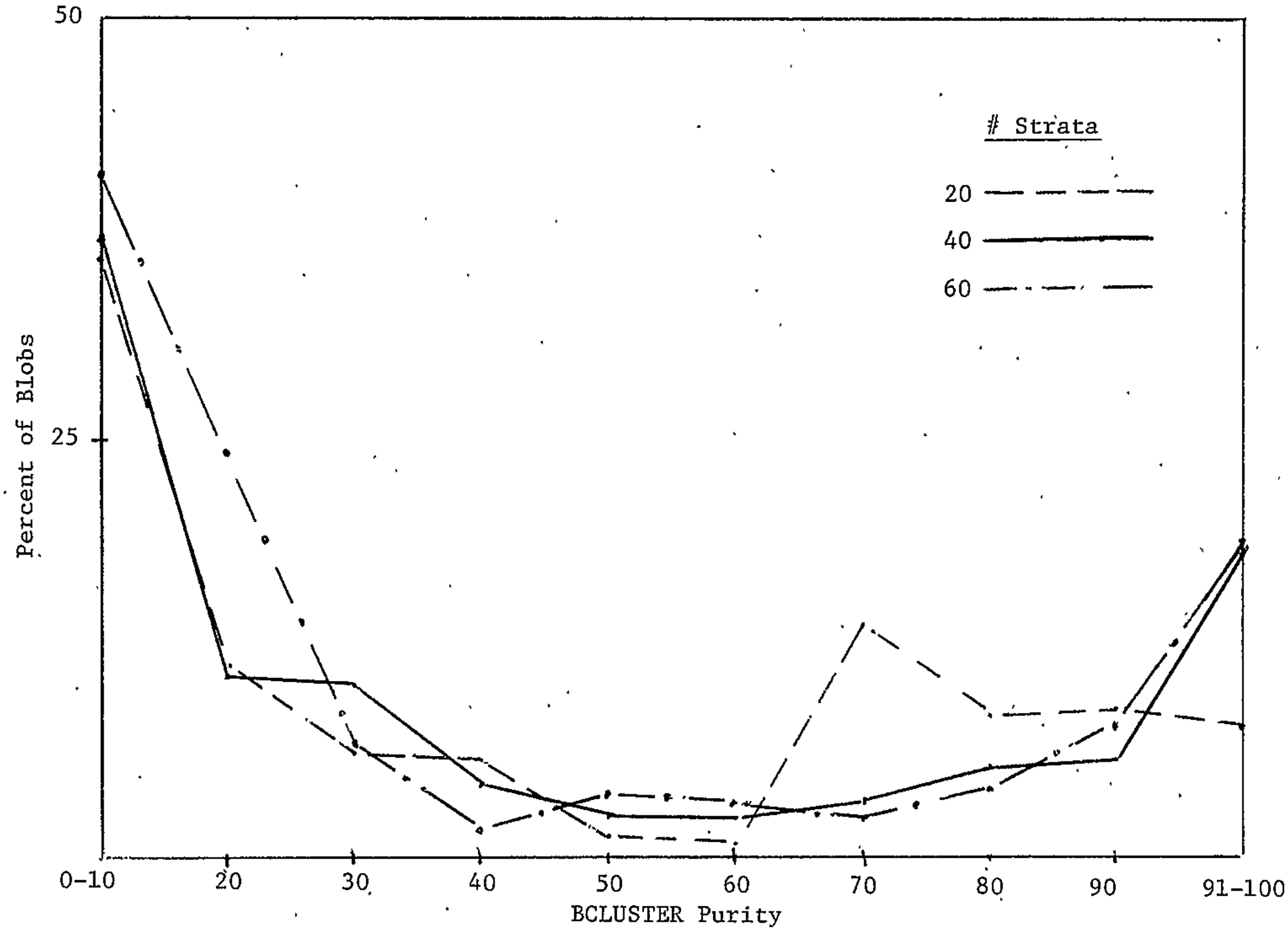


FIGURE 28. BCLUSTER PURITY FOR 13 NORTH DAKOTA TY78 SEGMENTS

7
PREPARATION OF A DATA BASE FOR SMALL GRAINS

A substantial data base was prepared in support of a number of planned research activities, including both development and testing of ideas and procedures. A general discussion of this data base follows, and specific details and tables are presented in Appendix L.

The data base consists of 65 5 x 6 nmi segments located throughout the United States Great Plains during the 1976 and 1977 winter and spring wheat growing seasons. Each segment consists of data from all available usable acquisitions (average 8 acquisitions) of Landsat data (Landsats 1 and 2), merged together with wall-to-wall ground truth inventory data supplied by USDA and prepared by JSC/LEC. Important pixel-by-pixel information generated during processing, such as cloud/water/shadow identifications or blob numbers, were retained with each segment.

In order to support the variety of tasks depending on this data base, segments were selected after initial screening into two categories, A and B. The principle differences between the two categories were in the segment selection procedures and the state to which the segments were processed. For some applications it is appropriate to use Category A segments for development and training, and Category B segments for test and evaluation.

The segment selection procedures used were as follows. First the 171 segments available to us were screened to eliminate those whose acquisition history, ground truth quality, and data quality are inadequate. When carried out, 107 segments remained for possible selection. These segments were stratified by APU* and year. Category B segments were then selected by randomly sampling a quota (1/3) of each stratum, subject to the constraint that at least one segment be selected from each stratum.

*Agrophysical Units, as used within LACIE.

By this procedure, 36 segments were selected to represent a nearly random samples of the variability of the region. The remaining segments were considered candidates for Category A. Stricter minimum limits on acquisition history and data quality were observed, so that selected segments were well suited for development work. However, the selections were more qualitative, taking into account desire for more acquisitions versus desire for a nice spread to cover variability. Even though selecting Category A segments first would have resulted in more and nicer segments for development work, it was necessary to select them second in order to maintain the statistical integrity of the Category B sample used to support test and evaluation.

In order to prepare the data for use and to minimize unwanted variability due to external effects, preprocessing steps were applied to all segments. Table 31 lists the algorithms used, and they are also detailed in Appendix L.

TABLE 31. PROCESSING STEPS APPLIED TO THE SMALL GRAINS DATA BASE

Function	Algorithm
Ground Truth Manipulation	(CONVRT, FLDS15)
Merge Acquisitions and Ground Truth	(MERGE)
Data Screening (clouds, water, etc)	(SCREEN)
Satellite Calibrations Cosine Sun Angle Spatially Varying Haze Correction	(XSTAR)
Feature Extraction	(TASCAP)
Field Pattern Identification*	(SUPERB, STRIP)
Extraction of Field Means*	(COMPRS)

*Category A segments only.

The Category A segments received additional processing steps designed to identify field patterns and obtain spectral mean values for those field patterns, both including and excluding pixels adjacent to the field pattern boundaries. The results of this processing are a data set reduced in size by a factor of roughly twenty, appropriate for easy and efficient repeated used by many development tasks.

The need for processing to identify field patterns is due to the lack of field information as part of the available ground truth. The identification is carried out by the algorithm SUPERB, which is described in Appendix D. In short, SUPERB is a spatial-spectral clustering algorithm similar to BLOB [44] which forms field-like patches, but which, unlike BLOB, can be supervised by the ground truth, so that only adjacent pixels that contain like ground truth codes are assigned to the same field pattern.

Since some tasks require the data in its non-compressed pixel form, that is provided as well. Codes for each pixel giving the field pattern number and field boundary indicator are included with the Category A pixel data.

To summarize, the data base consists of 67 5x6-mile segments selected from 171 LACIE Phase 2 and Phase 3 blind sites. The segments are divided into two groups: one was chosen with fine acquisition history and quality and was processed to produce field means; the other was chosen with a good statistical spread. Each segment contains from 3 to 15 Landsat acquisitions, preprocessed through a standard battery of algorithms (Table 31).. A segment consists of 23,000 pixels, each with between 31 and 91 channels that include standardized Landsat, screening, ground truth, field pattern, boundary and other data.

8
RECOMMENDATIONS

This section summarizes the major recommendations resulting from the reported investigations. Together with the Executive Summary, it forms a concise account of the year's effort and its ramifications.

At the broadest level, we recommend that an overall information system context be borne in mind to guide development of area estimation technology and its component techniques. One realization of this context should be through implementation of a baseline estimation system having, at least in embryo, all of the anticipated components or functions perceived for the final system. This context will provide the environment for a phased evolution of technology through definable and evaluatable stages.

Regarding overall considerations of objective labeling techniques, we recommend that:

- The second year of the planned effort to understand the physical foundations supporting objective labeling of wheat and other small grains be continued with the focus on near-term development of a labeling procedure based on such understanding.
- The same general approach be initiated toward understanding the problem of labeling corn and soybeans.

More specific recommendations are indicated below relative to several aspects of the objective labeling investigations:

- Refined machine labeler for spring wheat and barley:
  - that the accuracy of labeling be evaluated on training and test data sets;
  - that proportion estimation performance be evaluated within a Procedure M context, with both ground-truth and analyst labels, using the 18-segment TY 1978 data set.

- Understanding of current labeling technology:
  - that a follow-on large-scale analyst experiment be designed and conducted to further investigate analyst labeling performance;
  - that the following features be made a part of that experiment:
    - Multi-analyst (to evaluate the overall distribution of accuracy of a team of analysts).
    - Multicrop (to evaluate not only the precision of analysts labels within a crop, but to identify confusion crops).
    - Confidence labeling (to examine whether clear accuracy trends appear as a function of an analyst's confidence in a given label).
    - Independent and cooperative labeling (to evaluate comparative accuracy among independent analysts as compared to team approaches or labels fabricated from multiple analyst inputs, e.g., vote and average labels).
    - Multidate (to evaluate accuracy as a function of time of year and missing acquisitions).
    - Multitarget (to compare pixels, quasi-fields, and actual fields as labeling targets).
    - Multistratum (to evaluate analyst performance as a function of temporal-spectral strata defined to characterize crop spectral phenology).
- Feature definition:
  - that the Tasseled-Cap transformation be used and analysis of spectral data structures and their agronomic basis be continued and extended to the spectral space of the Thematic Mapper;
  - that relationships between reflectance space and Landsat space be confirmed and/or refined through expanded joint analysis of LACIE/LACIE-Transition field measurement and Landsat data.

- that development and use of a meteorologically driven model of the spectral phenology of wheat be continued, and that the availability of needed technology for extension to corn and soybeans be ascertained.

- Feature extraction:
    - that data normalization precede profile fitting and other information extraction operations to reduce variability due to factors other than crop type and crop condition;
    - that our recommended steps for other parts of profile fitting be followed (e.g., where appropriate, use non-linear fitting procedures and the second model form after crop calendar shift operations; with first model form, be sure to use day and Greenness offsets);
    - that development of temporal-spectral profile technology be continued for other crops and for additional levels of complexity of application, such as crop development stage estimation;
    - that improved aids be developed to assist analysts in labeling field-like targets, e.g., field-delineation overlays and spectral aids.

- Signature characterization:
    - that a signature extraction procedure employing CLASSY be tested on a small scale using blob means as elements;
    - that temporal-spectral profile parameters be analyzed as variables and characterized for crop and competing crop signatures.

- Procedure development:
  - that incorporation of temporal-spectral profile parameters in analyst-oriented labeling procedures be investigated;
  - that an analyst-assisted machine labeler for small grains be investigated;
  - that analyst-interpreters be investigated as extractors of collateral data for labeling and that other aspects of man-machine interactions be considered.

Regarding overall machine processing considerations, we recommend that:

- The stratified area estimation technology identified in Section 3 be utilized as a baseline crop acreage estimation environment for further technique development in the near future.
- Procedure M, because of its modularity, versatility, statistical framework and physical foundation, be considered as a candidate to represent SAE technology.

More specifically directed at machine processing component technology, we recommend that:

- Normalization techniques, like the Landsat 3 to Landsat 2 transform and others incorporated in Procedure M, be utilized to provide a standard frame of reference for the analysis of agronomic remote sensing data.
- Development of baseline normalization and feature extraction techniques for the thematic mapper be initiated using field measurement and/or simulated data.

- The inherent separability of Landsat data be evaluated using a non-parametric technique, like the nearest neighbor algorithm.
- The use of derived features like crop calendar shift, peak Greenness, and other that are agronomically interpretable be investigated as features for unsupervised stratification.
- A field-finding algorithm, like BLOB, AMOEBA or SUPERB, be utilized to distinguish pure and mixed pixels; and that comparative tests of unsupervised clustering algorithms be run on each group separately.
- Static spectral-temporal stratification be further explored to establish standard trajectory strata; and that physically and statistically based stratification strategies be compared.
- Study of optimal sample allocation techniques like Bayes sequential and Neyman be continued and, additionally, that work be initiated to:
    - evaluate their potential with static strata;
    - evaluate them in the presence of analyst error to avoid designing strategies that may be optimal with respect to ground truth, but yet magnify analyst-induced mean square error;
    - develop strategies that sample to minimize overall variance which is composed of expected sampling and labeling variances, and also include cost factors.
- Contextual estimation approaches using strata based on pixel composition (field center, mixed, trash, or misregistered) be developed.

- The nearest neighbor algorithm be evaluated as a classifier and as a mechanism to extend "high confidence" labels.
- Development be continued of a predictive model for stratified area estimation procedures based on segment data, establishing factors that affect estimates that are observable in imagery to parametrize the model.
- Development of information theoretic based performance measures be continued, including establishment of guidelines for interpreting these measures, especially in a multicrop, multisensor application.

REFERENCES

1. The LACIE Symposium: Proceedings of the Plenary Session (JSC-14551) and Independent Peer Evaluation of the Large Area Crop Inventory Experiment (JSC-14550), NASA Johnson Space Center, Houston, TX, October 1978.

2. The LACIE Symposium: "The USDA Applications Test System", Technical Sessions, Vol. II (JSC-16015), July 1979, pp. 1067-1125.

3. Cicone, R., E. Crist, R. Kauth, P. Lambeck, W. Malila, and W. Richardson, Development of Procedure M for Multicrop Inventory, With Tests of a Spring Wheat Configuration, Final Report 132400-16-F, Environmental Research Institute of Michigan, Ann Arbor, MI, March 1979.

4. Bizzell, R., A. Feiveson, F. Hall, M. Bauer, B. Davis, W. Malila, and D. Rice, Crop Identification Technology Assessment for Remote Sensing (CITARS), Vol. X: Interpretation of Results, JSC-09393, NASA Johnson Space Center, Houston, TX, December 1975.

5. Myers, R. M. and S. G. Wheeler, Fisher Linear Classification and Cross-Validation Bias Correction Comparison With the LACIE Procedure 1 Small Grains Estimation, IBM Report RES 23-70, International Business Machines Corporation, Federal Systems Division, Houston, TX, August 1979.

6. Peters, B. C., Jr. and H. F. Walker, "An Iterative Procedure for Obtaining Maximum Likelihood Estimates of the Parameters of a Normal Distribution", SIAM J. Appl. Math., Vol. 35, No. 2, September 1978, pp. 362-378.

7. Lennington, R. K. and M. E. Rassbach, Mathematical Description and Program Documentation for CLASSY, An Adaptive Maximum Likelihood Clustering Method, Technical Memorandum LEC-12177, Lockheed Electronics Co., April 1979.

8. Peters, Charles, "Spectral/Spatial Models for Agricultural Landsat Data", NAS-9-15543, Mod. 2S, Quarterly SR&T Review, September 11, 1979.

9. Kauth, R. J. and W. Richardson, Procedure B: A Multispectral Training Selection and Proportion Estimation Procedure for Processing Landsat Agricultural Data, Technical Report 122700-31-F, Environmental Research Institute of Michigan, Ann Arbor, MI, November 1977.

REFERENCES (Cont'd)

10. LACIE Phase III Accuracy Assessment, Final Report, LACIE-00478, JSC-13766, NASA Johnson Space Center, Houston, TX, August 1979.

11. Balon, R. J. and R. C. Cicone, Uniform Color Space Analysis of LACIE Image Products, Technical Memorandum 132400-10-R, Environmental Research Institute of Michigan, Ann Arbor, MI, May 1979.

12. Kauth, R. J. and G. S. Thomas, "The Tasseled-Cap -- A Graphic Description of the Spectral-Temporal Development of Agricultural Crops As Seen by Landsat", Machine Processing of Remotely Sensed Data, Symposium Proceedings, Purdue/LARS, West Lafayette, IN, 1976.

13. Kauth, R., P. Lambeck, W. Richardson, G. Thomas, and A. Pentland, "Feature Extraction Applied to Agricultural Crops as Seen by Landsat", Proceedings of the Technical Sessions, Vol. II, The LACIE Symposium, JSC 16015, NASA Johnson Space Center, Houston, TX, July 1979, pp. 705-721.

14. Pore, M. D. and R. A. Abotten, "A Programmed Labeling Approach to Image Interpretation", Proceedings of the Technical Sessions, Vol. II, The LACIE Symposium, NASA Johnson Space Center, Houston, TX, July 1979, pp. 923-936.

15. Thomson, D. R. and O. A. Wehmanen, "Application of Landsat Digital Data for Monitoring Drought", Proceedings of the Technical Sessions, Vol. I, The LACIE Symposium, July 1979, pp. 431-438.

16. Badhwar, G. D., "Small Grains" presented at NASA Supporting Research and Technology Quarterly Review, NASA Johnson Space Center, Houston, TX, September 1978.

17. Hay, C. M., R. W. Thomas, and A. S. Benson, Development of Techniques for Producing Static Strata Maps and Development of Photointerpretation Methods Based on Multitemporal Landsat Data; Final Report. NASA Contract NAS9-14565, Principal Investigator, R. N. Colwell. University of California, Space Sciences Laboratory, Series 19, Issue 1, Berkeley, CA, August 1977.

18. Malila, W. A. and E. P. Crist, "Spectral Separability of Spring Wheat from Other Spring Small Grains", presented at NASA SR&T Quarterly Review, NASA Johnson Space Center, Houston, TX, September 1978; (Documented in ERIM Report 132400-12-P).

REFERENCES (Cont'd)

19. Kauth, R. P. Lambeck, W. Richardson, G. Thomas, and A. Pentland, "Feature Extraction Applied to Agricultural Crops as Seen by Landsat", Proceedings of the Technical Sessions, Vol. II, The LACIE Symposium, JSC-16015, NASA Johnson Space Center, Houston, TX, July 1979, pp. 705-721.

20. Ritchie, J. T,, U.S.D.A., Grassland, Soil and Water Research Laboratory, Temple, TX (private communications, March and May 1979).

21. Suits, G. H., "The Calculation of the Directional Reflectance of a Vegetative Canopy", Remote Sensing of Environment, Vol. 2: 117-125, 1972.

22. Turner, R. E., Radiative Transfer in Real Atmospheres, Final Report 190100-24-T, Environmental Research Institute of Michigan, Ann Arbor, MI, July 1974.

23. Kaneko, T., Analysis of Methods of Generating Bias and Scale Factors for PFC Guns, IBM Memo RES 20-53-134, August 12, 1976.

24. Kraus, G., Another Method for Computing PFC Scaling Factors, NASA Memo TF3/76-630, September 9, 1976.

25. Hocutt, W. T., An Improved Method of Calculating Scale Factors and Bias for Color Film Generation, LEC-8301, June 1976.

26. Cate, R. B., D. E. Phinney, M. C. Kinsler, M. L. Sestak, T. Hodges, and J. Dishler, Color Analysis of Landsat Digital Data Using a Cubic Color Model Based on Relative Energies, Lockheed Electronics Co. (to be published in early 1980)

27. Cicone, R. and R. Balon, "Multicrop Labeling Aids", Quarterly Progress Report on Analysis of Scanner Data for Crop Inventories, ERIM 132400-12-P, Environmental Research Institute of Michigan, Ann Arbor, MI, June-September 1978.

28. Walker, M. M., "Principal Components Colour Display of ERTS Imagery", Proceedings of the Second Canadian Symposium on Remote Sensing, Guelph, Ontario, Canada, 1974.

29. Juday, Richard D., Colorimetric Principles as Applied to Multichannel Imagery, NASA Johnson Space Center, Houston, TX, Revision B, July 15, 1978.

REFERENCES (Cont'd)

30. Lambeck, P. F., Signature Extension Preprocessing for Landsat MSS Data, Final Report 122700-32-F, Environmental Research Institute of Michigan, Ann Arbor, MI, November 1977.

31. Boullion, T. L., B. S. Duran, and P. L. Odell, Estimation and Classification With Incomplete Data, Annual Report, University of Texas at Dallas, JSC-09703, June 1, 1974 - May 31, 1975.

32. Helmer, D. C., C. V. Nazare, R. W. Payne, and W. L. West, III, Report on the Analyst Test Using the Blob Labeling Procedures, Lockheed Electronics Corp., AD 63-1827-4845-33, May 17, 1979.

33. Wheeler, S. G., "Procedure 1A Development", presented at NASA SR&T Quarterly Review, NASA Johnson Space Center, Houston, TX, September 10, 1979.

34. Malila, W. A., R. C. Cicone, and J. M. Gleason, Wheat Signature Modeling and Analysis for Improved Training Statistics, Final Report 109600-66-F, Environmental Research Institute of Michigan, Ann Arbor, MI, May 1976.

35. Malila, W. A., J. M. Gleason, and R. C. Cicone, Atmospheric Modeling Related to Thematic Mapper Scan Geometry, Final Report 119300-5-F, Environmental Research Institute of Michigan, Ann Arbor, MI, April 1976.

36. Kauth, R. J., P. F. Lambeck, and G. S. Thomas, "SCREEN, A Procedure for Automatically Detecting Garbled Data, Clouds, Snow, Cloud Shadows, and Water in Landsat MSS Data", New Technology Report, Environmental Research Institute of Michigan, Ann Arbor, MI, November 10, 1977.

37. Lambeck, P. F., "Spatially Varying XSTAR Haze Correction", New Technology Report, Environmental Research Institute of Michigan, Ann Arbor, MI, March 30, 1979.

38. Kauth, R. J. and J. F. Hemdal, "Landsat 1 to Landsat 2 Data Conversion", New Technology Report, Environmental Research Institute of Michigan, Ann Arbor, MI, November 21, 1977.

39. Cochran, William G., Sampling Techniques, Third Edition, John Wiley & Sons, NY, 1977, pp. 89-90.

REFERENCES (Cont'd)

40. Richardson, W., A Study of the Tolerance Block Approach to Spectral Stratification, Technical Memorandum 132400-25-R, Environmental Research Institute of Michigan, Ann Arbor, MI, August 1979.

41. Lennington, K., "P1A Unsupervised Procedure Development", presented at NASA SR&T Quarterly Review, NASA Johnson Space Center, Houston, TX, LEC/SSD REF 642-7701, September 1979.

42. Kauth, R. J. and W. Richardson, "Task 1, Multisegment Training", Analysis of Scanner Data for Crop Inventories, Third Quarterly Report on NASA Contract NAS9-15476, ERIM Report 132400-12-P, Environmental Research Institute of Michigan, Ann Arbor, MI, September 1978, pp. 1-5.

43. Pont, W. Frank, "Appendix B: Effects of Random Stratification", from Technical Memorandum 132400-25-R, Environmental Research Institute of Michigan, Ann Arbor, MI, August 1979, pp. 49-55.

44. Kauth, R. J., A. P. Pentland, and G. S. Thomas, "BLOB, An Unsupervised Clustering Approach to Spatial Preprocessing of MSS Imagery", Eleventh International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan, Ann Arbor, MI, Vol. 2, 1977, pp. 1309-1317.

45. Kuretz, C. A. and C. M. Hay, "AI Procedures for Episodal Events", presented at NASA SR&T Quarterly Review, NASA Johnson Space Center, Houston, TX, June 1979.

46. Badhwar, G. D., "Multitemporal Classification Scheme for Corn", presented at NASA SR&T Quarterly Review, NASA Johnson Space Center, Houston, TX, June 1979.

47. International Mathematical & Statistical Libraries, Inc., Reference Manual, LIB-007, January 1979 (Revised).

48. Jackson, R. D. and P. Pinter, U.S. Water Conservation Laboratory, Phoenix, Arizona, private communication, March 1979.

49. Malila, W., J. Gleason, F. Sadowski, R. Cicone, and E. Crist, "Applications of Modeling to Analysis and Processing of Landsat Data", Proceedings of Twelfth International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan, Ann Arbor, MI, April 1978.

REFERENCES (Cont'd)

50. Nalepka, R. F., J. E. Colwell, and D. P. Rice, Forecasts of Winter Wheat and Production Using Landsat Data, Final Report 114800-38-F, Environmental Research Institute of Michigan, Ann Arbor, MI, December 1977.

51. Malila, W. A., J. M. Gleason, and R. C. Cicone, "Multispectral System Analysis Through Modeling and Simulation", Proceedings of Eleventh International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan, Ann Arbor, MI, April 25-29, 1977.

52. Rouse, J., R. Haas, J. Schell, and D. Deering, "Monitoring Vegetation Systems in the Great Plains", Third ERTS Symposium, NASA/SP-351, I:309-317, 1973.

53. Wyszecki, G. and W. S. Stiles, Color Science, John Wiley & Sons, New York, 1967.

54. Kaneko, Toyohisa and Sharon Tinkler, LACIE Film Study, Generation of Color Infrared Films, 78-FM-8, JSC-13847, NASA Johnson Space Center, Houston, TX, February 1978.

55. Coberly, W. A., Jack Tubbs, and P. L. Odell, The Influence of False Color Infrared Film on Training Errors in LACIE, University of Texas at Dallas, Annual Report, June 1, 1975 - August 31, 1976.

56. Austin, W. W., PFC Film Products for LACIE Segments, Memorandum to L. M. Flores, NASA, December 6, 1976.

57. Malila, W. A. and J. M. Gleason, Investigations of Spectral Separability of Small Grains, Early Season Wheat Detection, and Multicrop Inventory Planning, Final Report 122700-34-F, Environmental Research Institute of Michigan, Ann Arbor, MI, November 1977.

58. Anderson, T. W., An Introduction to Multivariate Statistical Analysis, John Wiley & Sons, N.Y., 1958.

59. Suits, G. H., "The Cause of Azimuthal Variations in Directional Reflectance of Vegetative Canopies", Remote Sensing of Environment, Vol. 2: 175-182, 1972.

REFERENCES (Cont'd)

60. Condit, H. R., "The Spectral Reflectance of American Soils", Photogrammetric Engineering, Vol. XXXVI, No. 9, September 1970, pp. 955-966.

61. Cover, T. M. and P. E. Hart, "Nearest Neighbor Pattern Classification", IEEE Trans. IT, Vol. IT-13, No. 1, January 1967.

INTENTIONALLY

LEFT

BLANK

DISTRIBUTION LIST

NASA Contract NAS9-15476

Name	Number of Copies
NASA/Johnson Space Center Earth Observations Division Houston, Texas 77058	
Attn: Dr. G.D. Badhwar/SF3	(1)
Attn: Mr. K. Baker/SF3	(1)
Attn: Mr. R.M. Bizzell/SF4	(1)
Attn: Mr. I.D. Browne/SF3	(2)
Attn: Mr. K. Bulow/SF6	(1)
Attn: Mr. K.J. Demel/SF3	(1)
Attn: Dr. J.W. Dietrich/SF3	(1)
Attn: Dr. J.E. Dornback/SF	(1)
Attn: Mr. J.L. Dragg/SF	(1)
Attn: Mr. R.B. Erb/SF	(1)
Attn: Dr. J.D. Erickson/SF3	(1)
Attn: Mr. J.G. Garcia/SF3	(1)
Attn: Mr. G.E. Graybeal/SF5	(1)
Attn: Dr. F. Hall/SF	(1)
Attn: Dr. C.R. Hallum/SF4	(1)
Attn: Mr. D.H. Hay/SF6	(1)
Attn: Mr. W.E. Hensley/SF2	(1)
Attn: Dr. R.P. Heydorn/SF3	(1)
Attn: Mr. R.O. Hill/SF4	(1)
Attn: Mr. W.C. Jones/SF3	(1)
Attn: Mr. R. Joosten/SF5	(1)
Attn: Mr. J.C. Lyon/SF6	(1)
Attn: Mr. T. Pendleton/SF3	(2)
Attn: Dr. D.E. Pitts/SF3	(1)
Attn: Research Data Facility/SF6	(2)
Attn: Dr. O.G. Smith/SF5	(1)
Attn: Mr. M.C. Trichel/SF3	(1)
Attn: Ms. Onalee Wald/SF2	(2)
Attn: Dr. F.P. Weber/SF5	(1)
NASA/Johnson Space Center Earth Observations Division Houston, Texas 77058	
Attn: Dr. B.L. Carroll/C09	(2)

DISTRIBUTION LIST (Cont.)

Name	Number of Copies
NASA/Johnson Space Center Technical Library Branch Houston, Texas 77058	
Attn: Ms. Retha Shirkey/JM6	(4)
NASA/Johnson Space Center Technical Support Procurement Branch Houston, Texas 77058	
Attn: Ms. E. McAnelly/BB63	(1)
NASA/Johnson Space Center Space and Life Sciences Directorate Houston, Texas 77058	
Attn: Mr. R.B. MacDonald	(1)
Attn: Mr. W.E. Rice	(1)
NASA/Ames Research Center Moffett Field, California 94035	
Attn: Dr. D.M. Deerwester	(1)
NASA/Goddard Space Flight Center Greenbelt, Maryland 20771	
Attn: Dr. J. Barker/923	(1)
Attn: Mr. W. Alford/563	(1)
NASA Headquarters Washington, DC 20546	
Attn: Mr. Pitt G. Thome/ER-2	(1)
Attn: Ms. Ruth I. Whitman/ERL-2	(1)
Agency for International Development Office of Science and Technology/SA-18 Washington, DC 20523	
Attn: Dr. Charles K. Paul	(1)

DISTRIBUTION LIST (Cont.)

Name	Number of Copies
Dr. G. O. Boatwright (EW/CCA) 1050 Bay Area Blvd.. Clear Lake City, Texas 77058	(1)
Duke University Department of Statistics 1804 Woodburn Durham, North Carolina 27705	
Attn: Dr. H.O. Hartley	(1)
Earth Resources Laboratory National Space Technology Laboratory Bay St. Louis, Mississippi 39520	
Attn: Mr. D.W. Mooneyhan	(1)
EROS Data Center U.S. Department of Interior Sioux Falls, South Dakota 57198	
Attn: Mr. G. Thorley	(1)
Attn: Mr. Allen Watkins	(1)
Environmental Protection Agency Western Environmental Research Laboratories P.O. Box 15027 Las Vegas, Nevada 89114	
Attn: Chief of Technical Support	(1)
IBM Corporation FSD, Mail Code 56 1322 Space Park Drive Houston, Texas 77058	
Attn: Mr. Stan Wheeler	(1)
NOAA Chief, Climatic Impact Assessment Division 3300 Whitehaven, NW Page Bldg. 2, Rm. 139 Washington, DC 20235	
Attn: Norton Strommen	(1)

DISTRIBUTION LIST (Cont.)

Name	Number of Copies
NOAA/NESS Office of Research Department of Commerce Washington, DC 20233	
Attn: Harold Yates	(1)
NOAA/Code OAX-1 Environmental Data and Information Services Washington, DC 20235	
Attn: Dr. Thomas Potter	(1)
NOAA CEAS Federal Building Columbia, Missouri 65201	
Attn: W. Wilson	(1)
Oregon State University Environmental Remote Sensing Applications Laboratory Corvallis, Oregon 97331	
Attn: Dr. Barry J. Schrumpf	(1)
Pennsylvania State University Space Science and Engineering Laboratory Office for Remote Sensing of Earth Resources 220 Electrical Engineering West Building University Park, Pennsylvania 16802	
Attn: Dr. Gary W. Peterson	(1)
Purdue University Purdue Industrial Research Park 1220 Potter Drive West Lafayette, Indiana 47906	
Attn: Dr. Marvin Bauer	(1)
Attn: Dr. David Landgrebe	(1)
Attn: Mr. Terry Phillips	(1)
Attn: Dr. Philip Swain	(1)

DISTRIBUTION LIST (Cont.)

Name	Number of Copies
South Dakota State University Director, Remote Sensing Institute Agriculture Engineering Building Brookings, South Dakota 57006	
Attn: Mr. Victor I. Myers	(1)
Texas A&M University Department of Mathematics College Station, Texas 77843	
Attn: Dr. L.F. Guseman, Jr.	(1)
Texas A&M University College of Agriculture Texas Agricultural Experiment Station Blackland Conservation Research Center P.O. Box 748 Temple, Texas 76501	
Attn: Dr. J.T. Ritchie	(1)
U.S. Department of Agriculture SEA/AR Bldg. 005, Rm. 125 BARC-West Beltsville, Maryland 20705	
Attn: Mr. Carl W. Carlson	(1)
U.S. Department of Agriculture SEA/NPS Bldg. 005, Rm. 233A BARC-West Beltsville, Maryland 20705	
Attn: Jerry Ritchie	(1)
U.S. Department of Agriculture Forest Service 240 W. Prospect St. Fort Collins, Colorado 80521	
Attn: Dr. Richard Driscoll	(1)

DISTRIBUTION LIST (Cont.)

Name	Number of Copies
U.S. Department of Agriculture FAS-CCAD 1050 Bay Area Blvd. Houston, Texas 77058	
Attn: Mr. Jimmy D. Murphy	(1)
U.S. Department of Agriculture Economics, Statistics, and Cooperatives Service Washington, DC 20250	
Attn: Mr. Rich Allen	(1)
Attn: Mr. Charles E. Caudill	(1)
Attn: Mr. Galen F. Hart	(1)
Attn: Mr. William E. Kibler	(1)
U.S. Department of Agriculture/SCS 1943 Newton Square East Reston, Virginia 22090	
Attn: Dr. R.H. Gilbert	(1)
U.S. Department of Agriculture Soil and Water Conservation Research Division P.O. Box 267 Weslaco, Texas 78596	
Attn: Dr. Craig Wiegand	(1)
U.S. Department of Interior Geological Survey 1925 Newton Square East Reston, Virginia 22070	
Attn: Mr. Fred Doyle	(1)
U.S. Department of Interior Geological Survey Federal Center Denver, Colorado 80225	
Attn: Dr. Harry W. Smedes	(1)

DISTRIBUTION LIST (Cont.)

Name	Number of Copies
University of Arkansas Mathematics Department Fayetteville, Arkansas 72704	
Attn: Dr. Jack D. Tubbs	(1)
University of California at Berkeley Forestry Department Berkeley, California 94720	
Attn: Robert Colwell	(1)
University of California at Berkeley 260 Space Sciences Lab Berkeley, California 94720	
Attn: Ms. Claire M. Hay	(2)
University of Texas at Dallas Box 688 Richardson, Texas 75080	
Attn: Dr. Patrick L. Odell	(1)
University of Tulsa Mathematics-Sciences Department 600 South College Tulsa, Oklahoma 74104	
Attn: Dr. W.A. Coberly	(1)

EARTH RESOURCES - NTIS